

STIC-Biotech/ChemLib

109561

**From:** Chan, Christina  
**Sent:** Wednesday, December 03, 2003 8:48 AM  
**To:** Swope, Sheridan; STIC-Biotech/ChemLib  
**Subject:** RE: 09830144

Please rush this one. Chris

*Chris Chan*

TC 1600 New Hire Training Coordinator and SPE 1644  
308-3973  
CM-1, 9B19

-----Original Message-----

**From:** Swope, Sheridan  
**Sent:** Tuesday, December 02, 2003 8:12 PM  
**To:** Chan, Christina  
**Subject:** FW: 09830144

I made a mistake on the search request.

-----Original Message-----

**From:** Swope, Sheridan  
**Sent:** Tuesday, December 02, 2003 8:08 PM  
**To:** Chan, Christina  
**Subject:** 09830144

Chris, May I have this rushed?--overdue amdt.

For 09/830,144, pls search and interference search:

**SID 1, nt 408-1091 against the NT and AA data bases.**

**SID 2, aa76-303 against the NT and AA data bases.**

**SID 3, nt1338-1541 against the NT and AA data bases.**

**SID 4, aa437-504 against the NT and AA data bases.**

Thanks!

Sheridan Swope, Ph.D.  
Patent Examiner, AU 1652  
Recombinant Enzymes  
sheridan.swope@uspto.gov  
703-305-1696 (voice)  
703-308-3014 (FAX)  
Mailbox: CM1 Rm10D01  
Office: CM1 Rm12D12

Point of Contact:  
Barb O'Bryen  
Technical Information Specialist  
STIC CM1 6A05 308-4291

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 11:40:09 ; Search time 137.149 Seconds  
(without alignments)  
1647.882 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504  
Perfect score: 359  
Sequence: 1 QSPULTLQSTWTHQSSSSS.....AEFYRLWSVDHGEQSVVTP 68

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+32n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09830144/runat\_03122003\_122511\_23420/app\_query.fasta\_1.654  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09830144 @cgn2\_1\_428 @runat\_03122003\_122511\_23420  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	359	100.0	1560	12	US-10-384-743-1	Sequence 1, Appli

2	359	100.0	1560	13	US-10-123-427-1	Sequence 1, Appli
3	359	100.0	1560	13	US-10-123-427-5	Sequence 5, Appli
4	359	100.0	1560	12	US-10-158-895-1	Sequence 1, Appli
5	359	100.0	1568	12	US-10-384-743-42	Sequence 42, Appli
6	359	100.0	1568	13	US-10-158-895-42	Sequence 42, Appli
7	359	100.0	1569	12	US-10-384-743-10	Sequence 10, Appli
8	359	100.0	1569	13	US-10-158-895-10	Sequence 10, Appli
9	359	100.0	16877	10	US-09-764-877-3349	Sequence 3349, Ap
10	352	98.1	696	10	US-09-925-300-330	Sequence 330, App
11	346	96.4	409	11	US-09-918-995-32946	Sequence 32946, A
12	75.5	21.0	672	12	US-09-928-267-16	Sequence 16, Appli
13	75.5	21.0	833	12	US-09-928-267-8	Sequence 8, Appli
14	75	20.9	1610	12	US-09-928-267-119	Sequence 119, App
15	71	19.8	996	10	US-09-974-300-6219	Sequence 6219, Ap
16	69.5	19.4	1492	10	US-09-941-831-7	Sequence 7, Appli
17	69.5	19.4	9507	11	US-09-764-891-9785	Sequence 9785, Ap
18	69	19.2	12409	11	US-09-989-442-156	Sequence 156, App
19	68.5	19.1	3068	14	US-10-259-453-1	Sequence 1, Appli
20	68.5	19.1	3082	14	US-10-259-453-2	Sequence 2, Appli
21	68.5	19.1	32767	12	US-10-004-113-4	Sequence 4, Appli
22	68	18.9	273	9	US-09-864-761-22569	Sequence 22569, A
23	68	18.9	8918	10	US-09-764-864-1682	Sequence 1682, Ap
24	68	18.9	8919	10	US-09-764-864-1683	Sequence 1683, Ap
25	67.5	18.8	2377	12	US-10-374-979-65	Sequence 65, Appli
26	67.5	18.8	2377	14	US-10-037-534-2	Sequence 2, Appli
27	67.5	18.8	2678	11	US-09-919-039-234	Sequence 234, App
28	66.5	18.5	743	12	US-10-027-632-147810	Sequence 147810,
29	66.5	18.5	743	12	US-10-027-632-147811	Sequence 147811,
30	66.5	18.5	743	13	US-10-027-632-147810	Sequence 147810,
31	66.5	18.5	743	13	US-10-027-632-147811	Sequence 147811,
32	66	18.4	684	12	US-10-029-386-24794	Sequence 24794, A
33	66	18.4	728	12	US-10-029-386-24125	Sequence 24125, A
34	66	18.4	902	12	US-10-027-632-31743	Sequence 31743, A
35	66	18.4	902	12	US-10-027-632-31744	Sequence 31744, A
36	66	18.4	902	13	US-10-027-632-31743	Sequence 31743, A
37	66	18.4	902	13	US-10-027-632-31744	Sequence 31744, A
38	66	18.4	1173	14	US-10-156-761-1415	Sequence 1415, Ap
39	66	18.4	9025608	14	US-10-156-761-1	Sequence 1, Appli
40	65.5	18.2	499	11	US-09-918-995-20193	Sequence 20193, A
41	65.5	18.2	4753	13	US-10-098-841-18	Sequence 18, Appli
42	65	18.1	895	10	US-09-764-891-5489	Sequence 5489, Ap
43	64.5	18.0	966	11	US-09-938-842A-2555	Sequence 2555, Ap
44	64.5	18.0	1409	14	US-10-012-542-82	Sequence 82, Appli
45	64.5	18.0	42432	12	US-10-029-120-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-384-743-1  
; Sequence 1, Application US/10384743  
; Publication No. US20030162228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JF98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1560  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

```
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
US-10-384-743-1

Alignment Scores:
Pred. No.: 1.16e-41 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x US-10-384-743-1 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAAGCCGACCTTAACCTGCAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 1398 TCTGACGGAGCGCTCTTCCGCTCCCGCCGCCACCTCGCTCCCGCTGGCGAGGACGGT 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTATGTGACTTTGCTGAGTTTACCGCTCTGGACCGTGGACCAATGGC 1517
QY 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 2
US-10-123-427-1
; Sequence 1, Application US/10123427
; Publication No. US20020119525A1
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/123.427
; FILING DATE: 17-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/406.854
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/752.891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..1541
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 30..1541
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-123-427-1

Alignment Scores:
Pred. No.: 1.16e-41 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x US-10-123-427-1 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAAGCCGACCTTAACCTGCAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 1398 TCTGACGGAGCGCTCTTCCGCTCCCGCCGCCACCTCGCTCCCGCTGGCGAGGACGGT 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTATGTGACTTTGCTGAGTTTACCGCTCTGGACCGTGGACCAATGGC 1517
QY 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 3
US-10-123-427-5
; Sequence 5, Application US/10123427
; Publication No. US20020119525A1
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/123.427
; FILING DATE: 17-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/406.854
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/752.891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
```

; FILING DATE: 24-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 17981/111  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202)672-5300  
 ; TELEFAX: (202)672-5399  
 ; TELEX: 904136  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1560 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 30..1541  
 ; FEATURE:  
 ; NAME/KEY: mat\_peptide  
 ; LOCATION: 30..1541  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
 US-10-123-427-5

Alignment Scores:  
 Pred. No.: 1,16e-41 Length: 1560  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-10-123-427-5 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1338 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGCAGCTCCAGC 1397  
 Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1398 TCTGACGAGGCGCTTTCCTGCTCCGCGCCGCCACTCGCTCCGCGCTGCGAGGACGGT 1457  
 Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1458 CGTGTGAGCCCTAATGAGCTTTGCTGAGTTTACCGCCCTCTGGAGCGTGACCATGGC 1517  
 Qy 61 GluGlnSerValValThrAlaPro 68  
 Db 1518 GAGCAGAGCGTGGTGACAGCACC 1541

RESULT 4  
 US-10-158-895-1  
 ; Sequence 1, Application US/10158895  
 ; Publication No. US20020155624A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ONO, KOICHIRO  
 ; APPLICANT: OHTOMO, TOSHIHIKO  
 ; APPLICANT: TSUCHIYA, MASAYUKI  
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
 ; FILE REFERENCE: 053466/0278  
 ; CURRENT APPLICATION NUMBER: US/10/158,895  
 ; CURRENT FILING DATE: 2002-06-03  
 ; PRIOR APPLICATION NUMBER: US/09/529,279  
 ; PRIOR FILING DATE: 2000-04-11  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
 ; PRIOR FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: JP 9/290188  
 ; PRIOR FILING DATE: 1997-10-22  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 1560  
 ; TYPE: DNA

; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (30)..(1541)  
 US-10-158-895-1  
 Alignment Scores:  
 Pred. No.: 1,16e-41 Length: 1560  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-10-158-895-1 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1338 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGCAGCTCCAGC 1397  
 Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1398 TCTGACGAGGCGCTTTCCTGCTCCGCGCCGCCACTCGCTCCGCGCTGCGAGGACGGT 1457  
 Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1458 CGTGTGAGCCCTAATGAGCTTTGCTGAGTTTACCGCCCTCTGGAGCGTGACCATGGC 1517  
 Qy 61 GluGlnSerValValThrAlaPro 68  
 Db 1518 GAGCAGAGCGTGGTGACAGCACC 1541

RESULT 5

US-10-384-743-42  
 ; Sequence 42, Application US/10384743  
 ; Publication No. US20030162228A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ONO, KOICHIRO  
 ; APPLICANT: OHTOMO, TOSHIHIKO  
 ; APPLICANT: TSUCHIYA, MASAYUKI  
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
 ; FILE REFERENCE: 053466/0278  
 ; CURRENT APPLICATION NUMBER: US/10/384,743  
 ; CURRENT FILING DATE: 2003-03-11  
 ; PRIOR APPLICATION NUMBER: US/09/529,279  
 ; PRIOR FILING DATE: 2000-04-11  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
 ; PRIOR FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: JP 9/290188  
 ; PRIOR FILING DATE: 1997-10-22  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 42  
 ; LENGTH: 1568  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (11)..(1549)  
 US-10-384-743-42

Alignment Scores:  
 Pred. No.: 1,17e-41 Length: 1568  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-10-384-743-42 (1-1568)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1338 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGCAGCTCCAGC 1397  
 Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1398 TCTGACGAGGCGCTTTCCTGCTCCGCGCCGCCACTCGCTCCGCGCTGCGAGGACGGT 1457  
 Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1458 CGTGTGAGCCCTAATGAGCTTTGCTGAGTTTACCGCCCTCTGGAGCGTGACCATGGC 1517  
 Qy 61 GluGlnSerValValThrAlaPro 68  
 Db 1518 GAGCAGAGCGTGGTGACAGCACC 1541

Db 1346 CAAAGCCGACCTTAACCTCGAGTCACCAACAGCAGCAGCAGCAGCAGCTCCAGC 1405  
QY 21 SerAspGlyClyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 1406 TCTGACGGAGGCTCTTCGCTCCCGCCCGCCACATCGCTCCCGCTGGCGAGGACGGT 1465  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
Db 1466 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 1525  
QY 61 GluGlnSerValValThrAlaPro 68  
Db 1526 GAGCAGAGCGTGTGACAGCACCG 1549

## RESULT 6

US-10-158-895-42  
; Sequence 42, Application US/10158895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin ver. 2.1  
; SEQ ID NO 42  
; LENGTH: 1568  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (11)..(1549)  
US-10-158-895-42

Alignment Scores:  
Pred. No.: 1,17e-41 Length: 1568  
Score: 359.00 Matches: 68  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-10-158-895-42 (1-1568)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 1346 CAAAGCCGACCTTAACCTCGAGTCACCAACAGCAGCAGCAGCAGCAGCTCCAGC 1405  
QY 21 SerAspGlyClyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 1406 TCTGACGGAGGCTCTTCGCTCCCGCCCGCCACATCGCTCCCGCTGGCGAGGACGGT 1465  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
Db 1466 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 1525  
QY 61 GluGlnSerValValThrAlaPro 68  
Db 1526 GAGCAGAGCGTGTGACAGCACCG 1549

## RESULT 7

US-10-384-743-10  
; Sequence 10, Application US/10384743  
; Publication No. US20030162228A1

; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; PRIOR FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 1569  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1557)  
US-10-384-743-10

Alignment Scores:  
Pred. No.: 1,17e-41 Length: 1569  
Score: 359.00 Matches: 68  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-10-384-743-10 (1-1569)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 1315 CAAAGCCGACCTTAACCTCGAGTCACCAACAGCAGCAGCAGCAGCAGCTCCAGC 1374  
QY 21 SerAspGlyClyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 1375 TCTGACGGAGGCTCTTCGCTCCCGCCCGCCACATCGCTCCCGCTGGCGAGGACGGT 1434  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
Db 1435 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 1494  
QY 61 GluGlnSerValValThrAlaPro 68  
Db 1495 GAGCAGAGCGTGTGACAGCACCG 1518

## RESULT 8

US-10-158-895-10  
; Sequence 10, Application US/10158895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 1569





Job time : 146.149 secs

Db 769 ACTCAGATGACTTCCTCAGCCAGCAGCCTGAGAGGCCCCGAGACGTGAAGCTGGCCCCGCA 828

Level of Agreement	Percentage of Respondents
Strongly agree	10%
Agree	25%
Disagree	15%
Strongly disagree	5%
Don't know	45%

44 ProTyrValAspPheAlaGluPheTyrArgLeuTrpSerVal---AspHisGly 60

Db 877 ---TACGTGCTGGTGGCTGGCCACTACCCCGTGTGTCCATAGCCGAGCACGGG 927

US-09-974-300-6219

Patent No. US20020146721A1

; APPLICANT: Clausen, Ib Groth

FILE OF INVENTION: EXPRESSION  
; FILE REFERENCE: 10085.500-US

; CURRENT FILING DATE: 2001-10-05  
 ; PRIOR APPLICATION NUMBER: 00/590 500

; PRIOR APPLICATION NUMBER: 60/279,526

```

; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

;
; LENGTH: 996
; TYPE: DNA

```

US-09-974-300-6219

Alignment scores:	
Pred. No.:	1.67
Length:	996

Percent Similarity: 50.00%  
Conservative: 13

DB:	10	0	3
Gaps:	10	0	3

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-09-974-300-6219 (1-996)

[illegible]

26 PheArgSerArgPro---A]aHisSerLeuProProG]vG]uAsnG]vAArgA]C]uP

Dbb  
134 TTTATGAACGAACCGACGTCGTTTCAGAACCTCCTGACTACGATCCAAGGAAACGTC

27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054

On 55 v[er]g[e]l m[u]lti-v[er]g[e]l - m[u]lti-v[er]g[e]l

Dbb 233 ATTACAGATCCA 244

Search completed: December 4, 2003 14:31:41

Search completed: December 4, 2003, 14:31:41

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 09:43:02 : Search time 27.3378 Seconds  
(without alignments)  
1097.894 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSPRLTQSTNTHQSSSS.....AEFYRLMSVDHGEQSVVTAP 68

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 599978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09830144/runat\_03122003\_122510\_23346/app\_query.fasta\_1.654  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144@cgn 1\_76 @runat\_03122003\_122510\_23346 -NCPUS=6  
-NO MMAP -LARGSEQURY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359	100.0	1560	2	US-08-752-891-1
2	359	100.0	1560	2	US-08-752-891-5
3	359	100.0	1560	2	US-09-144-178-1
4	359	100.0	1560	2	US-09-144-178-5
5	359	100.0	1560	3	US-09-406-854-1
6	359	100.0	1560	3	US-09-406-854-5
7	359	100.0	1560	4	US-09-529-279-1
8	359	100.0	1560	4	US-10-158-895-1
9	359	100.0	1568	4	US-09-529-279-42
10	359	100.0	1568	4	US-10-158-895-42
11	359	100.0	1569	4	US-09-529-279-10
12	359	100.0	1569	4	US-10-158-895-10

c 13	71.5	19.9	1108	5	PCT-US93-03035-1	Sequence 1, Appli
14	67.5	18.8	2559	2	US-09-070-060-2	Sequence 2, Appli
15	67.5	18.8	2559	3	US-09-357-746-2	Sequence 2, Appli
16	67.5	18.8	5359	2	US-09-070-060-1	Sequence 1, Appli
17	67.5	18.8	5372	3	US-09-357-746-1	Sequence 1, Appli
c 18	65	18.1	951	4	US-09-252-991A-7708	Sequence 7708, Ap
19	65	18.1	1008	4	US-09-252-991A-7923	Sequence 7923, Ap
20	65	18.1	1227	4	US-09-252-991A-7853	Sequence 7853, Ap
c 21	64.5	18.0	1409	4	US-09-461-325-82	Sequence 82, Appl
22	64.5	18.0	2259	4	US-09-252-991A-6446	Sequence 6446, Ap
c 23	64	17.8	564	4	US-09-252-991A-15454	Sequence 15454, A
24	64	17.8	2319	4	US-09-252-991A-15284	Sequence 15284, A
c 25	64	17.8	2427	4	US-09-252-991A-15396	Sequence 15396, A
c 26	63.5	17.7	1101	4	US-09-152-060-23	Sequence 23, Appl
27	63.5	17.7	1173	3	US-08-706-216-5	Sequence 5, Appli
28	63.5	17.7	1473	4	US-09-152-060-43	Sequence 43, Appl
c 29	63.5	17.7	2784	3	US-08-857-076-39	Sequence 39, Appl
c 30	63.5	17.7	3017	3	US-08-857-076-52	Sequence 52, Appl
c 31	63.5	17.7	3119	3	US-08-857-076-53	Sequence 53, Appl
c 32	62	17.3	336	4	US-09-252-991A-1069	Sequence 1069, Ap
c 33	62	17.3	1395	2	US-08-553-367A-1	Sequence 1, Appli
c 34	62	17.3	1395	3	US-09-295-306-1	Sequence 1, Appli
c 35	62	17.3	1395	4	US-09-734-719-1	Sequence 1, Appli
c 36	62	17.3	1575	4	US-09-252-991A-1139	Sequence 1139, Ap
c 37	62	17.3	2859	4	US-09-252-991A-944	Sequence 944, Ap
c 38	62	17.3	5977	3	US-09-024-020B-1	Sequence 1, Appli
39	62	17.3	5977	4	US-09-425-043-1	Sequence 1, Appli
40	62	17.3	6007	3	US-09-024-020B-2	Sequence 2, Appli
41	62	17.3	6007	4	US-09-425-043-2	Sequence 2, Appli
42	62	17.3	6556	3	US-09-024-020B-7	Sequence 7, Appli
43	62	17.3	6556	4	US-09-425-043-7	Sequence 7, Appli
44	62	17.3	6586	3	US-09-024-020B-43	Sequence 43, Appl
45	62	17.3	6586	4	US-09-425-043-43	Sequence 43, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-752-891-1  
; Sequence 1, Application US/08752891  
; Patent No. 5837819  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1560 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 30..1541  
NAME/KEY: mat\_peptide  
LOCATION: 30..1541  
US-08-752-891-1

Alignment Scores:  
Pred. No.: 9.72e-37 Length: 1560  
Score: 359.00 Matches: 68  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-08-752-891-1 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 1338 CAAAGCCGACCTTAACCTCGAGTCACCAACACGACGACGAGCAGCTCCAGC 1397  
QY 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 1398 TCTGACGAGCGCTTTCGCTCCGCGCCGCCACCTCGCTCCGCGCGAGGACGGT 1457  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
Db 1458 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGACGCTGGACCAATGCC 1517  
QY 61 GluGlnSerValValThralaPro 68  
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

## RESULT 2

US-08-752-891-5  
Sequence 5, Application US/08752891  
Patent No. 5837819  
GENERAL INFORMATION:  
APPLICANT: MATSUMOTO, Kunihiro  
APPLICANT: NISHIDA, Eisuke  
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/752,891  
FILING DATE: 20-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-126282

FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1560 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 30..1541  
NAME/KEY: mat\_peptide  
LOCATION: 30..1541  
US-08-752-891-5  
Alignment Scores:  
Pred. No.: 9.72e-37 Length: 1560  
Score: 359.00 Matches: 68  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0  
US-09-830-144-4\_COPY\_437\_504 (1-68) x US-08-752-891-5 (1-1560)  
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 1338 CAAAGCCGACCTTAACCTCGAGTCACCAACACGACGACGAGCAGCTCCAGC 1397  
QY 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 1398 TCTGACGAGCGCTTTCGCTCCGCGCCGCCACCTCGCTCCGCGCGAGGACGGT 1457  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
Db 1458 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGACGCTGGACCAATGCC 1517  
QY 61 GluGlnSerValValThralaPro 68  
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541  
RESULT 3  
US-09-144-178-1  
Sequence 1, Application US/09144178  
Patent No. 5989862  
GENERAL INFORMATION:  
APPLICANT: MATSUMOTO, Kunihiro  
APPLICANT: NISHIDA, Eisuke  
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/144,178  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/752,891  
FILING DATE: 20-NOV-1996  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1560 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 30..1541  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 30..1541  
US-09-144-178-1

Alignment Scores:  
Pred. No.: 9,72e-37 Length: 1560  
Score: 359.00 Matches: 68  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-09-144-178-1 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 1338 CAAGCCGACCTTAACCTTCAGTCCACACGACGACGACGACGACGACGACGACG 1397  
Qy 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 1398 TCTGACGAGGACCTTTCCTCCGCGCCGACCTCCGCTCCGCTCCGCTCCGCTCC 1457  
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCCTCTGGAGCGTGACCATGGC 1517  
Qy 61 GluGlnSerValValThrAlaPro 68  
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 4  
US-09-144-178-5  
Sequence 5, Application US/09144178  
Patent No. 5989862  
GENERAL INFORMATION:  
APPLICANT: MATSUMOTO, Kunihiro  
APPLICANT: NISHIDA, Eisuke  
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/144,178  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/752,891  
FILING DATE: 20-NOV-1996  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1560 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 30..1541  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 30..1541  
US-09-144-178-5

Alignment Scores:  
Pred. No.: 9,72e-37 Length: 1560  
Score: 359.00 Matches: 68  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-09-144-178-5 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 1338 CAAGCCGACCTTAACCTTCAGTCCACACGACGACGACGACGACGACGACGACG 1397  
Qy 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 1398 TCTGACGAGGACCTTTCCTCCGCGCCGACCTCCGCTCCGCTCCGCTCCGCTCC 1457  
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCCTCTGGAGCGTGACCATGGC 1517  
Qy 61 GluGlnSerValValThrAlaPro 68  
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 5  
US-09-406-854-1  
Sequence 1, Application US/09406854  
Patent No. 6140042  
GENERAL INFORMATION:  
APPLICANT: MATSUMOTO, Kunihiro  
APPLICANT: NISHIDA, Eisuke  
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR

NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Review #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/406,854  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/752,891  
FILING DATE: 20-NOV-1996  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1560 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 30..1541  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 30..1541  
US-09-406-854-1

Alignment Scores:		
Pred. No.:	9, 72e-37	Length: 1560
Score:	359.00	Matches: 68
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	3	Gaps: 0

US-09-830-144-4 COPY 437 504 (1-68) x US-09-406-854-1 (1-1560)

Qy	1	GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer	20
Db	1338	CAAAAGCCGACCTTAAACCTTGACCTCCACCAACAGCACGACGACGACGCTCCAGC	1397
Qy	21	SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly	40
Db	1398	CTCGACGGAGGACCTCTCCGCTCCCGCCGCCCCACCTCGCTCCCGCTGGCAGGACGGT	1457
Qy	41	ArgValGluProTrpValAspPheAlaGluPheTrpArgLeuTrpSerValAspHisGly	60
Db	1458	CGTGTTCAGCCCATGTGGACTTTCCTCAGTTTACCGCCTCTGGACGCTGGACCATGGC	1517
Qy	61	GluGlnSerValValThrAlaPro	68
Db	1518	GAGCAGAGCCGTGGTGACAGCACCG	1541

RESULT 6  
US-09-406-854-5  
; Sequence 5, Application US/09406854  
; Patent No. 6140042  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/406,854  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1560 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 30..1541  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 30..1541  
; US-09-406-854-5

Alignment Scores:		
Pred. No.:	9.72e-37	Length:
Score:	359.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	100.00%	Indels:
DB:	3	Gaps:

US-09-830-144-4 COPY 437 504 (1-68) X US-09-406-854-5 (1-1560)

Qy	1	GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db	1338	CAAAAGCCGACCTTAACCTTCGAGTCACCAACACGACACGAGACGACGCTCCAGC 1397
Qy	21	SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db	1398	TCCTGACGAGGAGCCTCTTCGCTCCGCGCCGCGCCACTCGCTCCGCTCGCAGAGACGGT 1457

```
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
DB 1458 CGTGTGAGCCCTATGTGGACCTTCTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 1517
QY 61 GluGlnSerValValThrAlaPro 68
DB 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 7
US-09-529-279-1
; Sequence 1, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
US-10-158-895-1
Alignment Scores:
Pred. No.: 9,72e-37 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x US-10-158-895-1 (1-1560)
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
DB 1338 CAAAGCCCGACCTTAACCTTCAGTCCACCAACACGACGACGAGCAGCAGCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
DB 1398 TCTGACGAGCGCTCTTCCTCCGCTCCGCCGCCACTCGCTCCGCCCTGGAGCGTGGACCATGGC 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
DB 1458 CGTGTGAGCCCTATGTGGACCTTCTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 1517
QY 61 GluGlnSerValValThrAlaPro 68
DB 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 8
US-10-158-895-1
; Sequence 1, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
```

```
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
US-10-158-895-1
Alignment Scores:
Pred. No.: 9,72e-37 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x US-10-158-895-1 (1-1560)
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
DB 1338 CAAAGCCCGACCTTAACCTTCAGTCCACCAACACGACGACGAGCAGCAGCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
DB 1398 TCTGACGAGCGCTCTTCCTCCGCTCCGCCGCCACTCGCTCCGCCCTGGAGCGTGGACCATGGC 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
DB 1458 CGTGTGAGCCCTATGTGGACCTTCTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 1517
QY 61 GluGlnSerValValThrAlaPro 68
DB 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 9
US-09-529-279-42
; Sequence 42, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(1549)
US-09-529-279-42
Alignment Scores:
Pred. No.: 9,79e-37 Length: 1568
Score: 359.00 Matches: 68
```

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-09-529-279-42 (1-1568)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1346 CAAGCCGACCTTAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1405  
 QY 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1406 TCTGACGAGGCGCTTCCGCTCCCGCCGCCACTCGCTCCCGCTCGGAGGACGGT 1465  
 QY 41 ArgValGluProTrpValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 Db 1466 CGTGTGAGCCCTATGTGACCTTTCGTTTACCGCTCTGGAGCGTGGACCAATGGC 1525  
 QY 61 GluGlnSerValValThrAlaPro 68  
 Db 1526 GAGCAGAGCGTGTGACAGCAGCG 1549

RESULT 10

US-10-158-895-42  
 ; Sequence 42, Application US/10158895  
 ; Patent No. 6551840  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ONO, KOICHIRO  
 ; APPLICANT: OHTOMO, TOSHIHIKO  
 ; APPLICANT: TSUCHIYA, MASAYUKI  
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
 ; FILE REFERENCE: 053466/0278  
 ; CURRENT APPLICATION NUMBER: US/10/158,895  
 ; CURRENT FILING DATE: 2002-06-03  
 ; PRIOR APPLICATION NUMBER: US/09/529,279  
 ; PRIOR FILING DATE: 2000-04-11  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
 ; PRIOR FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: JP 9/290188  
 ; PRIOR FILING DATE: 1997-10-22  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 42  
 ; LENGTH: 1568  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (11)..(1549)  
 US-10-158-895-42

Alignment Scores:  
 Pred. No.: 9,796-37 Length: 1568  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-10-158-895-42 (1-1568)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1346 CAAGCCGACCTTAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1405  
 QY 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1406 TCTGACGAGGCGCTTCCGCTCCCGCCGCCACTCGCTCCCGCTCGGAGGACGGT 1465  
 QY 41 ArgValGluProTrpValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 Db 1466 CGTGTGAGCCCTATGTGACCTTTCGTTTACCGCTCTGGAGCGTGGACCAATGGC 1525

QY 61 GluGlnSerValValThrAlaPro 68  
 Db 1526 GAGCAGAGCGTGTGACAGCAGCG 1549

RESULT 11

US-09-529-279-10  
 ; Sequence 10, Application US/09529279  
 ; Patent No. 6451617  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ONO, KOICHIRO  
 ; APPLICANT: OHTOMO, TOSHIHIKO  
 ; APPLICANT: TSUCHIYA, MASAYUKI  
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
 ; FILE REFERENCE: 053466/0278  
 ; CURRENT APPLICATION NUMBER: US/09/529,279  
 ; CURRENT FILING DATE: 2000-04-11  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
 ; PRIOR FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: JP 9/290188  
 ; PRIOR FILING DATE: 1997-10-22  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 10  
 ; LENGTH: 1569  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (7)..(1557)  
 US-09-529-279-10

Alignment Scores:  
 Pred. No.: 9,8e-37 Length: 1569  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-09-529-279-10 (1-1569)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1315 CAAGCCGACCTTAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1374  
 QY 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1375 TCTGACGAGGCGCTTCCGCTCCCGCCGCCACTCGCTCCCGCTCGGAGGACGGT 1434  
 QY 41 ArgValGluProTrpValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 Db 1435 CGTGTGAGCCCTATGTGAGGACTTTCGTTTACCGCTCTGGAGCGTGGACCAATGGC 1494  
 QY 61 GluGlnSerValValThrAlaPro 68  
 Db 1495 GAGCAGAGCGTGTGACAGCAGCG 1518

RESULT 12  
 US-10-158-895-10  
 ; Sequence 10, Application US/10158895  
 ; Patent No. 6551840  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ONO, KOICHIRO  
 ; APPLICANT: OHTOMO, TOSHIHIKO  
 ; APPLICANT: TSUCHIYA, MASAYUKI  
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
 ; FILE REFERENCE: 053466/0278  
 ; CURRENT APPLICATION NUMBER: US/10/158,895  
 ; CURRENT FILING DATE: 2002-06-03  
 ; PRIOR APPLICATION NUMBER: US/09/529,279  
 ; PRIOR FILING DATE: 2000-04-11  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796

; PRIOR FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: JP 9/290188  
 ; PRIOR FILING DATE: 1997-10-22  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 10

; LENGTH: 1569  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (7)..(1557)  
 US-10-158-895-10

Alignment Scores:  
 Pred. No.: 9, 8e-37 Length: 1569  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x US-10-158-895-10 (1-1569)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1315 CAAAGCCGAGCTTAACCTGCAGTCCACCAACACGACACGACGAGCAGCTCCAGC 1374  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1375 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCCGCTCCGCGGAGGACGGT 1434  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1435 CGTCTGAGCCCTATGAGACTTCTGAGTTCCTACCGCCCTCTGAGCGCTGACCATGGC 1494  
 QY 61 GluGlnSerValValThrAlaPro 68  
 Db 1495 GAGCAGAGCGTGGTGACAGCACCG 1518

RESULT 13  
 PCT-US93-03035-1/c  
 ; Sequence 1, Application PC/TUS9303035  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ABBOTT LABORATORIES  
 ; TITLE OF INVENTION: PURIFIED THERMOSTABLE ENDONUCLEASE  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ABBOTT LABORATORIES  
 ; STREET: D-377 AP6D, ONE ABBOTT PARK ROAD  
 ; CITY: ABBOTT PARK  
 ; STATE: ILLINOIS  
 ; COUNTRY: USA  
 ; ZIP: 60064-3500  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US93/03035  
 ; FILING DATE: 19930330  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/869,306  
 ; FILING DATE: 16-APR-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/860,702  
 ; FILING DATE: 31-MAR-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BRAINARD, THOMAS D  
 ; REGISTRATION NUMBER: 32,459  
 ; REFERENCE/DOCKET NUMBER: 5145.PC.01

; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 708-937-4884  
 ; TELEFAX: 708-937-2623  
 ; TELEX:

; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1108 base pairs  
 ; TYPE: NUCLEIC ACID  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: circular  
 ; MOLECULE TYPE: DNA (genomic)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 114..926  
 PCT-US93-03035-1

Alignment Scores:  
 Pred. No.: 4, 82 Length: 1108  
 Score: 71.50 Matches: 15  
 Percent Similarity: 70.37% Conservative: 4  
 Best Local Similarity: 55.56% Mismatches: 7  
 Query Match: 19.92% Indels: 1  
 DB: 5 Gaps: 1

US-09-830-144-4\_COPY\_437\_504 (1-68) x PCT-US93-03035-1 (1-1108)  
 QY 18 SerSerSerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGly 37  
 Db 203 AGCGTGAGCGCGAGCGGCGTCTCTCCACGCCCGCCCGCCCTTTTCCCGC 144  
 QY 38 GluAspGlyArgValGluPro 44  
 Db 143 ---GATGAAAGGTGAACCC 126

RESULT 14  
 US-09-070-060-2  
 ; Sequence 2, Application US/09070060  
 ; Patent No. 5976849  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hustad, Carolyn M.  
 ; APPLICANT: Ghildyal, Namit  
 ; TITLE OF INVENTION: Human E3 Ubiquitin Protein  
 ; TITLE OF INVENTION: Ligase  
 ; NUMBER OF SEQUENCES: 11  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: ZENECA Pharmaceuticals, Inc.  
 ; STREET: 1800 Concord Pike  
 ; CITY: Wilmington  
 ; STATE: DE  
 ; COUNTRY: USA  
 ; ZIP: 19850-5437  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/070,060  
 ; FILING DATE: 30-APR-1998  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/073,839  
 ; FILING DATE: 05-FEB-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Higgins, Patrick H  
 ; REGISTRATION NUMBER: 39,709  
 ; REFERENCE/DOCKET NUMBER: PHM.70312  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 302.886.4889  
 ; TELEFAX: 302.886.8221  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:

33	QY	-----SerLeuProProGlyGluuAspClyArgValGluProTyrValaspPhe	48
814	DB	GTAACCAAGCTCCCTGGCCACCTGTTGGCAGACAGATGGAC-----	858
49	QY	AlaGluPheTyrArgLeuTrpSerValaspHisGlyGluGln	62
959	DB	---CAGCAGCGGGGAGCTTTACTATGTAGATCATGTTTCAGAAA	997

Search completed: December 4, 2003, 12:45:02  
Job time : 31.3378 secs

```
; LENGTH: 2559 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-070-060-2

Alignment Scores:
Pred. No.:      47.8          Length:      2559
Score:           67.50       Matches:        23
Percent Similarity: 47.30%   Conservative: 12
Best Local Similarity: 31.08% Mismatches:    20
Query Match:     18.80%     Indels:       19
DB:              2         Gaps:          3

US-09-830-144-4_COPY_437_504 (1-68) x US-09-070-060-2 (1-2559)

QY             2 SerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 21
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db            634 TCTAGTCAGCGCTCTCTCCGCCGCACAAATACAATAACAATACATCTGNAGGAGCACCA 753

QY             22 AspGly-----GlyLeuPheArgSerArgProAlahis---- 32
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db            754 TCTGGATTAAATATTCCTTCTTTACTATATCTGGAGGCTCAGGCCCTTAGGCCATTAAATCT 813

QY             33 -----SerLeuProProGluAspGlyArgValGluProTyTrValAspPhe 48
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db            814 GPAATCTAAGCTCCCTTGCCACCCTGTGGGAGCAGAGATGGAC----- 858

QY             49 AlaGluPheTyArgLeuTrpSerValAsphisGlyGluGln 62
                ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db            859 ----CAGCAGCGCGGAGTTTACTATGTAGATCATGTCAGAAA 897
```

```

RESULT 15
US-09-357-746-2
; Sequence 2, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM.70312.N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-357-746-2

```

```

Alignment Scores:
Pred. No.:          47.8      Length:        2559
Score:             67.50     Matches:       23
Percent Similarity: 47.30%    Conservative:   12
Best Local Similarity: 31.08% Mismatches:    20
Query Match:      18.80%     Indels:       19
DB:                3         Gaps:           3

US-09-830-144-4_COPY_437_504 (1-68) x US-09-357-746-2 (1-2559)

Qy      2 SerProThrLeuThrGlnSerHisThrAsnThrHisThrGlnSerSerSerSerSer 21
              ||| | :||| |::| ::| ::| ::| ::| ::| ::| ::| ::| 
Db      694 TCTAGTACAGGGTCCTCTCGCCGCCACAATAACAATCAAAATCATCTGAAGGAGAACA 753
                               -----glyLeuPheArgSerArgProAlaHis--- 32
Qy      22 AspGly-----
              ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| 
Db      754 TCTGGATTAAATTTCTCCTACTATATCTTGAGGCCTCAGGCCCTTAGCCATTAAATCT 813
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 09:07:25 ; Search time 121.703 Seconds  
(without alignments)  
2090.550 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091

Perfect score: 1252

Sequence: 1 gttagagcttgcgcagttatc.....cattacagatccttgcag 684

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 1368560

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -US09830144/runat 03122003 122347 21456/app query.fasta\_1.1230  
-Q=/cgn2\_1/USFTO.spool/US09830144/runat 03122003 122347 21456/app query.fasta\_1.1230  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPCUI=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09830144 @CGN\_1\_78 @runat 03122003 122347 21456  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
-----					

1	1252	100.0	518	15	US-10-283-023-2	Sequence 2, Appli
2	1252	100.0	579	12	US-10-384-743-4	Sequence 4, Appli
3	1252	100.0	579	14	US-10-158-895-4	Sequence 4, Appli
4	1252	100.0	590	12	US-10-384-743-15	Sequence 15, Appli
5	1252	100.0	590	14	US-10-158-895-15	Sequence 15, Appli
6	371	29.6	349	15	US-10-106-698-6345	Sequence 6345, Ap
7	371	29.6	455	9	US-09-757-982-5	Sequence 5, Appli
8	371	29.6	455	12	US-10-094-749-2477	Sequence 2477, Ap
9	363.5	29.0	394	10	US-09-862-027-19	Sequence 19, Appl
10	355	28.4	1036	12	US-10-354-358-24	Sequence 24, Appl
11	355	28.4	1036	14	US-10-014-882-2	Sequence 2, Appli
12	353.5	28.2	746	15	US-10-153-668-436	Sequence 436, App
13	353.5	28.2	859	15	US-10-153-668-324	Sequence 324, App
14	353.5	28.2	892	15	US-10-153-668-438	Sequence 438, App
15	349.5	27.9	1097	12	US-10-288-798-12	Sequence 12, Appl
16	346	27.6	328	10	US-09-862-027-18	Sequence 18, Appl
17	345.5	27.6	847	14	US-10-143-133-2	Sequence 2, Appli
18	341	27.2	252	12	US-09-976-782-41	Sequence 41, Appl
19	341	27.2	254	12	US-09-976-782-30	Sequence 30, Appl
20	341	27.2	256	12	US-09-863-776-41	Sequence 41, Appl
21	338.5	27.0	835	10	US-09-947-199-2	Sequence 2, Appli
22	335.5	26.8	835	10	US-09-947-199-8	Sequence 8, Appli
23	331.5	26.5	966	10	US-09-771-161A-197	Sequence 197, App
24	326.5	26.1	256	12	US-09-976-782-40	Sequence 40, Appl
25	326.5	26.1	257	12	US-09-976-782-29	Sequence 29, Appl
26	323	25.8	257	11	US-09-823-187-46	Sequence 46, Appl
27	323	25.8	257	12	US-09-863-776-42	Sequence 42, Appl
28	323	25.8	821	12	US-10-171-404A-48	Sequence 48, Appl
29	320	25.6	263	10	US-09-840-704-5	Sequence 5, Appli
30	313.5	25.0	850	10	US-09-904-389-2	Sequence 2, Appli
31	312	24.9	265	9	US-09-797-039-10	Sequence 10, Appl
32	312	24.9	265	12	US-10-170-789-10	Sequence 10, Appl
33	312	24.9	277	9	US-09-815-915-13	Sequence 13, Appl
34	312	24.9	277	10	US-09-882-166-4	Sequence 4, Appli
35	312	24.9	277	12	US-10-393-316-13	Sequence 13, Appl
36	312	24.9	277	12	US-10-170-789-17	Sequence 17, Appl
37	312	24.9	277	15	US-10-172-088-6	Sequence 6, Appli
38	312	24.9	278	9	US-09-797-039-13	Sequence 13, Appl
39	312	24.9	278	9	US-09-922-138-18	Sequence 18, Appl
40	312	24.9	278	9	US-09-922-138-27	Sequence 27, Appl
41	312	24.9	278	9	US-09-910-150-17	Sequence 17, Appl
42	312	24.9	278	9	US-09-910-150-31	Sequence 31, Appl
43	312	24.9	278	10	US-09-842-582-4	Sequence 4, Appli
44	312	24.9	278	10	US-09-934-406-4	Sequence 4, Appli
45	312	24.9	278	12	US-10-170-789-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-10-283-023-2  
; Sequence 2, Application US/10283023  
; Publication No. US20030091573A1  
; GENERAL INFORMATION:  
; APPLICANT: Carroll, Joseph M.  
; TITLE OF INVENTION: Methods and compositions for the  
; diagnosis and treatment of hematological disorders using  
; FILE REFERENCE: MPI01-239PIRM  
; TITLE OF INVENTION: 16319  
; CURRENT APPLICATION NUMBER: US/10/283,023  
; CURRENT FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: 60/335,044  
; PRIOR FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-283-023-2

Alignment Scores: 4.12e-125 Length: 518  
Pred. No.: 518

Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 15 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-283-023-2 (1-518)

QY 1 GTAGAGCTTCGGCAGCTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGAGCC 60  
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyAla 95  
QY 61 TCGTTGAATCCAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115  
QY 121 CTGCATGTGTGAACCATCTCTAATATTGTAAAGCTTTATGAGCC 180  
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
QY 181 CAGTCTCCAGAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 240  
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArg 155  
QY 241 GACCTGAAACACCAACTTACTCTGCTGCTGAGGGGAGCAGTCTTAAATAATTTGAT 300  
Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
QY 421 GATTTATTCTTTGGGAAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 216 GlyLeuLeuTrpGluValLeuThrArgGlySerAsnTyrSerGluLysCysAspValPheSerTrp 235  
QY 481 GCTTTCCGAATCATGTGGGCTGCTTCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLeuLysAsnLeu 255  
QY 541 CCTAAGCCCATTCAGAGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGGAGAAATTTGTGAAAATAATGACTCACTTGTGCGGTACTTCCAGGAGCAGAT 660  
Db 276 SerMetGluGluLeuValLysLeuMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295

QY 661 GAGCCATTACAGTATCTTGTCTCAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303

RESULT 2  
US-10-384-743-4  
; Sequence 4, Application US/10384743  
; Publication No. US20030162228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOI CHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; PRIOR FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-384-743-4 (1-579)

QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGAGCC 60  
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyAla 95  
QY 61 TCGTTGAATCCAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115  
QY 121 CTGCATGTGTGAACCATCTCTAATATTGTAAAGCTTTATGAGCC 180  
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
QY 181 CAGTCTCCAGAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 240  
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArg 155  
QY 241 GACCTGAAACACCAACTTACTCTGCTGCTGAGGGGAGCAGTCTTAAATAATTTGAT 300  
Db 156 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysLeuLysCysAsp 175  
QY 301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATAACAGGGGAGTGTCTGTGG 360  
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195  
QY 361 ATGGCACCCTGAAGTTTTTCAAGGTAGTAAATTCAGTGAATAATGTGACGTCTTCAAGCTGG 420  
Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
QY 421 GATTTATTCTTTGGGAAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
Db 216 GlyLeuLeuTrpGluValLeuThrArgGlySerAsnTyrSerGluLysCysAspValPheSerTrp 235  
QY 481 GCTTTCCGAATCATGTGGGCTGCTTCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLeuLysAsnLeu 255  
QY 541 CCTAAGCCCATTCAGAGCCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGGAGAAATTTGTGAAAATAATGACTCACTTGTGCGGTACTTCCAGGAGCAGAT 660  
Db 276 SerMetGluGluLeuValLysLeuMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295

QY 661 GAGCCATTACAGTATCTTGTCTCAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303

RESULT 3  
US-10-158-895-4  
; Sequence 4, Application US/10158895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOI CHIRO

```

; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-4

Alignment Scores:
Pred. No.: 4,23e-125 Length: 579
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-830-144-1_COPY_408_1091 (1-684) x US-10-158-895-4 (1-579)
QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGCC 60
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyGlyAla 95
QY 61 TGCTTGATCCAGCTGCTCTTGATGGAATATCTCAAGGGGCTCTTTATATAATGTG 120
Db 96 CysLeuAsnProValCysLeuValMetGluTyAlaGluGlyGlySerLeuTyAsnVal 115
QY 121 CTGATGCTGTCGACCATTCACATGATCTGCTCCACGCAATGAGTTGGTGTGTTA 180
Db 116 LeuHisGlyAlaGluProLeuProTyTyThrAlaAlaHisAlaMetSerTrpCysLeu 135
QY 181 CAGTGTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 240
Db 136 GlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProLysAlaLeuIleHisArg 155
QY 241 GACCTGAAACCAACCACTTACTCTGCTGAGTGAATTAATGAGTGAATTAATGAGT 300
Db 156 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175
QY 301 TTTGGTACAGCTGTGACATTCACACATGACCAATTAACCAAGGGAGTGTGCTGG 360
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195
QY 361 ATGGCACTGAAGTTTTCAGGTAGTAAATTAAGTGAATAAATGTCAGCTGG 420
Db 196 MetAlaProGluValPheGluGlySerAsnTySerGluLysCysAspValPheSerTrp 215
QY 421 GGTATTATCTTTGGGAAGTGATACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCA 480
Db 216 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 235
QY 481 GCTTTCCGAATCATGTGGGTGTTCAATAAGTGTACTGACACACACATGACCACTGATAAATAATTA 540
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255
QY 541 CTTAAGCCCATTCAGAGCTTGATCTGTTGTGTGTTAAGATCTCTCCAGCGCCCT 600
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275
QY 601 TCAATGGAGAAATTTGAAATAATGACTCATCTGATCGGCTACTTCCAGGAGCAGAT 660
Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyPheProGlyAlaAsp 295

```

```

QY 661 GAGCCATTACAGTATCCTTTGTTCAG 684
Db 296 GluProLeuGlnTyProCysGln 303

RESULT 4
US-10-384-743-15
; Sequence 15, Application US/10384743
; Publication No. US20030162228A1
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/384,743
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-384-743-15

Alignment Scores:
Pred. No.: 4,25e-125 Length: 590
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-830-144-1_COPY_408_1091 (1-684) x US-10-384-743-15 (1-590)
QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGCC 60
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyGlyAla 95
QY 61 TGCTTGATCCAGCTGCTCTTGATGGAATATCTCAAGGGGCTCTTTATATAATGTG 120
Db 96 CysLeuAsnProValCysLeuValMetGluTyAlaGluGlyGlySerLeuTyAsnVal 115
QY 121 CTGATGCTGTCGACCATTCACATGATCTGCTCCACGCAATGAGTTGGTGTGTTA 180
Db 116 LeuHisGlyAlaGluProLeuProTyTyThrAlaAlaHisAlaMetSerTrpCysLeu 135
QY 181 CAGTGTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 240
Db 136 GlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProLysAlaLeuIleHisArg 155
QY 241 GACCTGAAACCAACCACTTACTCTGCTGAGTGAATTAATGAGTGAATTAATGAGT 300
Db 156 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175
QY 301 TTTGGTACAGCTGTGACATTCACACATGACCAATTAACCAAGGGAGTGTGCTGG 360
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195
QY 361 ATGGCACTGAAGTTTTCAGGTAGTAAATTAAGTGAATAAATGTCAGCTGG 420
Db 196 MetAlaProGluValPheGluGlySerAsnTySerGluLysCysAspValPheSerTrp 215
QY 421 GGTATTATCTTTGGGAAGTGATACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCA 480
Db 216 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 235

```

QY 481 GCTTTCCGAATCATGTGGCTCTTCATATAGTGTACTCGACCACTGATATAAAATTTA 540  
Db 236 AlaPheArgIleMetTriaAlaValHisAsnGlyThrArgProProLeuLeuLeuLeuLeu 255  
QY 541 CTAAGCCCATTCAGAGCCTGATGCTCGTGTCTTAAAGATCCTTCCAGGCGCCT 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGAGCAAAATGTGAAATAATGACTCACTTGCATCGGTACTTCCAGGAGCAGAT 660  
Db 276 SerMetGluGluLeuValHisLeuMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
QY 661 GAGCCATTACAGTATCCTTGTCTAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303  
RESULT 5  
US-10-158-895-15  
; Sequence 15, Application US/10158895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-895-15  
Alignment Scores:  
Pred. No.: 4,25e-125 Length: 590  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-158-895-15 (1-590)  
QY 1 GTACAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 60  
Db 76 ValGluLeuArgGlnLeuSerArgValHisHisProAsnIleValLysLeuTyrGlyAla 95  
QY 61 TGCTTGAATCCAGTGTCTTGTCTGATGGAATATGCTGAAGGGGCTCTTATATAATGTG 120  
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115  
QY 121 CTGCATGGTGTGAACCATTCGCATATATATCTGCTGCCACGCAATGAGTTGTGTTTA 180  
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerIleProCysLeu 135  
QY 181 CAGTGTTCACCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGG 240  
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155  
QY 241 GACCTGAACCAACCAACTTACTGCTGTGTCAGGGGGGACAGTCTTAAAAATTGTGAT 300  
Db 156 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
QY 301 TTGTGACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGCTGCTGG 360

Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTTP 195  
QY 361 ATGGCACCTGAAGTATTTTGAAGTAGTAAATACAGTGAAGAAATGACGCTTTCAGCTG 420  
Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTTP 215  
QY 421 GCTATATATCTTGGAGAGTATAACGCTCGGAAACCCCTTTGATGAGATTTGTTGCCCCA 480  
Db 216 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
QY 481 GCTTTCCGAATCATGTGGCTCTTCATATAGTGTACTCGACCACTGATATAAAATTTA 540  
Db 236 AlaPheArgIleMetTriaAlaValHisAsnGlyThrArgProProLeuLeuLeuLeu 255  
QY 541 CTAAGCCCATTCAGAGCCTGATGCTCGTGTCTTAAAGATCCTTCCAGGCGCCT 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGAGCAAAATGTGAAATAATGACTCACTTGCATCGGTACTTCCAGGAGCAGAT 660  
Db 276 SerMetGluGluLeuValHisLeuMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
QY 661 GAGCCATTACAGTATCCTTGTCTAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303  
RESULT 6  
US-10-106-698-6345  
; Sequence 6345, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides  
; FILE REFERENCE: PA005PI  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 6345  
; LENGTH: 349  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (340)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-10-106-698-6345  
Alignment Scores:  
Pred. No.: 8,23e-31 Length: 349  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 15 Gaps: 8  
US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-106-698-6345 (1-349)  
QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCTGC 63  
Db 109 GluAlaGluIleLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 128  
QY 64 TTGTAATCCA-----GTGTGCTTGTGATGAATATGCTGAAGGGGGCTCTTATATAAT 117  
Db 129 LeuGluProProAsnTyrGlyIleValThrGluIyrAlaSerLeuGlySerLeuTyrAsp 148

118	GTGCTGCATGGT-----GCTGAACCAATGGCCOATATTATATCTGCTGCCACGCAATGAGT	171
QY		
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		
129		
130		
131		
132		
133		
134		
135		
136		
137		
138		
139		
140		
141		
142		
143		
144		
145		
146		
147		
148		
149		
150		
151		
152		
153		
154		
155		
156		
157		
158		
159		
160		
161		
162		
163		
164		
165		
166		
167		
168		
169		
170		
171		
172		
173		
174		
175		
176		
177		
178		
179		
180		
181		
182		
183		
184		
185		
186		
187		
188		
189		
190		
191		
192		
193		
194		
195		
196		
197		
198		
199		
200		
201		
202		
203		
204		
205		
206		
207		
208		
209		
210		
211		
212		
213		
214		
215		
216		
217		
218		
219		
220		
221		
222		
223		
224		
225		
226		
227		
228		
229		
230		
231		
232		
233		
234		
235		
236		
237		
238		
239		
240		
241		
242		
243		
244		
245		
246		
247		
248		
249		
250		
251		
252		
253		
254		
255		
256		
257		
258		
259		
260		
261		
262		
263		
264		
265		
266		
267		
268		
269		
270		
271		
272		
273		
274		
275		
276		
277		
278		
27		

[illegible]

```

US-09-757-982-5
; Sequence 5, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; CURRENT FILING DATE: 2004-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-982-5

Alignment Scores:
Pred. No.:      8.78e-31      Length:      455
Score:          371.00      Matches:      81
Percent Similarity: 56.2%      Conservative: 40
Best Local Similarity: 37.6%      Mismatches:   80
Query Match:      29.63%      Indels:       14
DB:               9          Gaps:         8

US-09-830-144-1_COPY_408_1091 (1-684) x US-09-757-982-5 (1-455)

QY      4   GAGCTTCGCAGCTTATCCGGTGTGAACCATCTTATATTGTAAGCTTTATGAGCCTGC 63
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      53   GluAlaGluLeuSerValLeuSerHisArgAsnIleGlnPheTyrGlyValle 72

QY      64   TTGAATCCA-----GTGTGTCCTTGATGGAATATGCTGAAGGGGGCTCTTTATATAAT 117
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      73   LeuGluProProAsnTyrGlylleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92

```

;  
ORGANISM: Homo sapiens  
US-10-094-749-2477

Alignment Scores:  
Pred. No.: 8,786-31 Length: 455  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
Gaps: 8  
DB:

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-094-749-2477 (1-455)

```

QY 4 GAGCTTCGGCAGTTATCCGCTGTAACCATCTATATTTGAAAGCTTTATGAGCGCTGC 63
Db 53 GUAAlaGluLeuSerValLeuSerHisArgAsnIleGlnPheTyrGlyValIle 72
QY 64 TTGAATCCA-----GTGTGCTTGTGATGAATATGCTGAAGGGGCTTTATATAAT 117
Db 73 LeuGluProProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92
QY 118 GTGCTGCATGGT-----GCTGAACCATTCGCATATATATCTGCTGCCACCGAATGAGT 171
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109
QY 172 TGTGTTTACAGTTTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCACGCTA 231
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129
QY 232 ATTCAGAGGACCTGAAACCACTTACTGCTGTTGAGGGGGGACAGTTCTTAAAA 291
Db 130 IleHisArgAspLeuLysSerArgAsnValValIleAlaAlaAspGly---ValLeuLys 148
QY 292 ATTGTGATTTTGGTACAGCTGTGACATTCAGACACACATACCAATACAAAG----- 345
Db 149 IleCysAspPheCly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167
QY 346 GAGAGTGTCTGGTGGACCTGAACTTTTGAAGTAGTAAATACAGTGAATAATGT 405
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187
QY 406 GAGCTTTCAGCTGGGTATTTCTTTGGAGTGTATACCGCTCGAAACCTTTGAT 465
Db 188 AspThrTyrSerTyrGlyValValLeuTrpGluMetLeuThrArgGluValProPheLys 207
QY 466 GAGATGGTGGCCAGCTTTCCGAATCATCTGG---GCTCTTCATAATGCTACGACCA 522
Db 208 GlyLeuGluGly-----LeuGlnValAlaTrpLeuValValGluLysAsnGluArgLeu 225
QY 523 CCACGTGATAAAATTTACCTAAGCCCATTCAGAGCTGTATGACTGCTTGTGCTTAA 582
Db 226 ThrIleProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245
QY 583 GATCCTTCCAGCGCCCTTCAATCGAGGAATTCGAAATAATG 627
Db 246 AspAlaLysArgProSerPheGlnIleLeuSerIleLeu 260

```

RESULT 9

US-09-862-027-19  
Sequence 19, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:  
APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: NO. US20020142428A1 Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
CURRENT FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: US 09/345,473  
PRIOR FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19  
LENGTH: 394

;  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-862-027-19

Alignment Scores:  
Pred. No.: 5,396-30 Length: 394  
Score: 363.50 Matches: 83  
Percent Similarity: 54.09% Conservative: 36  
Best Local Similarity: 37.73% Mismatches: 84  
Query Match: 29.03% Indels: 17  
Gaps: 6  
DB:

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-862-027-19 (1-394)

```

QY 4 GAGCTTCGGCAGTTATCCGCTGTAACCATCTATATTTGAAAGCTTTATGAGCGCTGC 63
Db 50 GUAAlaLysLeuPheAlaMetLeuLysHisProAsnIleIleAlaLeuArgGlyValCys 69
QY 64 TTGAATCCA-----GTGTGCTTGTGATGAATATGCTGAAGGGGCTTTATATAAT 117
Db 70 LeuLysGluProAsnLeuCysLeuValMetGluPheAlaArgGlyGlyProLeuAsnArg 89
QY 118 GTGCTGCATGGTGAACCATTCGCATATATATCTGCTGCCACCGAATGAGTGGTGT 177
Db 90 ValLeuSerGlyLysArgIleProAspIle-----LeuValAsnTrpAla 105
QY 178 TTACAGTGTTCCTCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATC 237
Db 106 ValGlnIleAlaArgGlyMetAsnTyrLeuHisAspGluAlaIleValProIleIleHis 125
QY 238 AGGACCTGAAACCCACCACTTACTGCTGGT-----GCAGGG 276
Db 126 ArgAspLeuLysSerSerAsnIleLeuIleLeuGlnLysValGluAsnGlyAspLeuSer 145
QY 277 GGGACAGTCTTAAATTTGTGATTTGTACAGCTGTGAC---ATTGACAGACACATG 333
Db 146 AsnLysIleLeuLysIleThrAspPheGlyLeuAlaArgGluTrpHisArgThrLys 165
QY 334 ACCAATAACAAGGGAGTGTCTGCTGGATGGCACCTGAAAGCTTTTGAAGTAGTAAATAC 393
Db 166 MetSerAlaAlaGlyThrTyrAlaTrpMetAlaProGluValIleArgAlaSerMetPhe 185
QY 394 AGTGAATAATGACGCTTTCAGCTGGGTATTTATTTCTTTGGAGTGAATGACGCTCG 453
Db 186 SerLysGlySerAspValTrpSerTyrGlyValLeuLeuTrpGluLeuLeuThrGlyGlu 205
QY 454 AAACCTTTTCATGAGATTTGGTGGCCAGCTTTCCGAATCATCTGGGCTGT---CAT 510
Db 206 ValProPheArgGlyIleAspGly-----LeuArgValAlaTyrGlyValAlaMetAsn 223
QY 511 GGTACTCGACCACTGATAAAAAATTTACCTAAGCCCATTCAGAGCTGTATGACTGCT 570
Db 224 LysLeuAlaLeuProIleProSerThrCysProGluProPheAlaLysLeuMetGluAsp 243
QY 571 TGTGCTTAAAGATCCTTCCCGCGCCCTTCAATGAGGAATTTGAAATAATGACT 630
Db 244 CysTrpAsnProAspProHisSerArgProSerPheThrAsnIleLeuAspGlnLeuThr 263

```

RESULT 10

US-10-354-358-24  
Sequence 24, Application US/10354358  
Publication No. US20030157082A1  
GENERAL INFORMATION:  
APPLICANT: Millennium Pharmaceuticals, Inc  
APPLICANT: Hunter, John Joseph  
APPLICANT: MacBeth, Kyle J.  
APPLICANT: Tsai, Fong-Ying  
APPLICANT: Lescon, Andrea  
APPLICANT: Lightcap, Eric S.  
APPLICANT: Williamson, Mark  
APPLICANT: Rudolph-Owen, Laura A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,

```
; TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,
; TITLE OF INVENTION: 3703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,
; TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,
; TITLE OF INVENTION: 8990, 2100, 9288, 6498, 10480, 20893, 33230, 1586, 9943,
; TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,
; TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES
; FILE REFERENCE: MPI02-020PIRNONIM
; CURRENT APPLICATION NUMBER: US/10/354,358
; CURRENT FILING DATE: 2003-01-30
; PRIOR APPLICATION NUMBER: US 60/353,600
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/364,517
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/371,075
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: US 60/371,507
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: US 60/372,984
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/374,194
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/382,995
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/385,023
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: US 60/388,853
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: US 60/389,395
; PRIOR FILING DATE: 2002-06-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-354-358-24

Alignment Scores:
Pred. No.: 5,55e-29 Length: 1036
Score: 355.00 Matches: 83
Percent Similarity: 53.48% Conservative: 40
Best Local Similarity: 36.09% Mismatches: 85
Query Match: 28.35% Indels: 22
DB: 12 Gaps: 6

US-09-830-144-1_COPY_408_1091 (1-684) x US-10-354-358-24 (1-1036)
QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGGAGCGCTGC 63
Db 171 GluAlaArgLeuPheAlaMetLeuArgHisProAsnIleIleGluLeuArgGlyValCys 190
QY 64 TTGAATCCA-----GTGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db 191 LeuGlnGlnProHisLeuCysLeuValLeuPheAlaArgGlyGlyAlaLeuAsnArg 210
QY 118 GTGCTGCATGCTGCTGAACCATTTGCCATATTACTGCT----- 156
Db 211 AlaLeuAlaAlaAlaAlaProAspProArgAlaProGlyProArgAlaArg 230
QY 157 -----GCCACGCAATG---AGTTGGTGTATTACAGTTGTTCCCAAGGAGTGGCTTAT 204
Db 231 ArgileProProHisValLeuValAsnTrpAlaValGlnIleAlaArgGlyMetLeuTrp 250
QY 205 CTTACAGCATGCAACCCAAAGCGTAATTCACAGGAGCTGAACCAACCAACTTACTG 264
Db 251 LeuHisGluGluAlaPheValProIleLeuHisArgAspLeuLysSerSerAsnIleLeu 270
QY 265 CTGCTGTCAGGG-----GGACAGCTTCTAAATTTGTGATTTT 303
Db 271 LeuLeuGluLysIleGluHisAspAspIleCysAsnLysThrLeuLysIleThrAspPhe 290
QY 304 GGTACAGCTGTGAC---ATTACAGACACATGACCAATAACAAGGGAGTGTGCTTGG 360
```

```
Db 291 GlyLeuAlaArgGluTrpHisArgThrThrLysMetSerThrAlaGlyThrTyAlaTrp 310
QY 361 ATGGCACCTGAAGTTTTTGAAGGTAGTAATATACAGTGAATAATATGACGCTCTTACGCTG 420
Db 311 MetAlaProGluValIleLysSerSerLeuPheSerLysGlySerAspIleTrpSerTrp 330
QY 421 GGTATTATCTTTGGGAAGTGATTAACCGGTGGAAACCCCTTTGATGAGATTTGGTGGCCCA 480
Db 331 GlyValLeuLeuTrpGluLeuLeuThrGlyGluValProTyrArgGlyIleAspGlyLeu 350
QY 481 GCTTTCCGAATCATGTGGCTGTTCAATAATGGTACTGCACCAACCATCATGATAAAAAATTA 540
Db 351 AlaValAlaTyrGlyValAlaValAsnLysLeuThrLeu---ProIleProSerThrCys 369
QY 541 CCTAAGGCCCATTTGAGAGCTGATGACTGCTGTTGGTCTAAGATCCCTTCCAGCGCCCT 600
Db 370 ProGluProPheAlaLysLeuMetLysGluCysTrpGlnGlnAspProHisIleArgPro 389
QY 601 TCAATGGAGGAATTTGAAATAATGACT 630
Db 390 SerPheAlaLeuIleLeuGluGlnLeuThr 399

RESULT 11
US-10-014-882-2
; Sequence 2, Application US/10014882
; Publication No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Kie, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding ti
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,744
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1036)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-014-882-2

Alignment Scores:
Pred. No.: 5,55e-29 Length: 1036
Score: 355.00 Matches: 83
Percent Similarity: 53.48% Conservative: 40
Best Local Similarity: 36.09% Mismatches: 85
Query Match: 28.35% Indels: 22
DB: 14 Gaps: 6

US-09-830-144-1_COPY_408_1091 (1-684) x US-10-014-882-2 (1-1036)
QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGGAGCGCTGC 63
Db 171 GluAlaArgLeuPheAlaMetLeuArgHisProAsnIleIleGluLeuArgGlyValCys 190
QY 64 TTGAATCCA-----GTGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db 191 LeuGlnGlnProHisLeuCysLeuValLeuPheAlaArgGlyGlyAlaLeuAsnArg 210
QY 118 GTGCTGCATGCTGCTGAACCATTTGCCATATTACTGCT----- 156
Db 211 AlaLeuAlaAlaAlaAlaProAspProArgAlaProGlyProArgAlaArg 230
QY 157 -----GCCACGCAATG---AGTTGGTGTATTACAGTTGTTCCCAAGGAGTGGCTTAT 204
```

Db 231 ArgileProHisValLeuValAsnTrpAlaValGlnIleAlaArgGlyMetLeuTyr 250  
 QY 205 CTTACAGCATGCAACCCAAAGCGTAATTCACAGGACCTGAAACCCAACTTACTG 264  
 Db 251 LeuHisGluGluAlaPheValProIleLeuHisArgAspLeuLysSerAsnIleLeu 270  
 QY 265 CTGGTTGACGGG-----GGACAGTTCTCTAAATAATTGTCATTTT 303  
 Db 271 LeuLeuGluLysIleGluHisAspAspIleCysAsnLysThrLeuLysIleThrAspPhe 290  
 QY 304 GGTACAGCTGTGAC---ATTACAGACACATGACCAATAACAAAGGGAGTCTGCTGG 360  
 Db 291 GlyLeuAlaArgGluTrpHisArgThrThrLysMetSerThrAlaGlyThrTrpAlaTrp 310  
 QY 361 ATGGCACCTGAAGTTTGAAGTACTAATTCAGTGAATAATGACCTCTTCAGCTGG 420  
 Db 311 MetAlaProGluValIleLysSerSerLeuPheSerLysGlySerAspIleTrpSerTyr 330  
 QY 421 GGTATTATCTTTGGGAAGTGATACCGCTCGGAACCCCTTTGATGAGATTGTTGCCCA 480  
 Db 331 GlyValLeuLeuTrpGluLeuLeuThrGlyGluValProTyrArgGlyIleAspGlyLeu 350  
 QY 481 GCTTTCCGAATCATGTGGCTGTTTCAATGCTGCTGACACCACTGATAAAATTTA 540  
 Db 351 AlaValAlaTyrGlyValAlaValAsnLysLeuThrLeu---ProIleProSerThrCys 369  
 QY 541 CCTAACCCCATTCAGAGCCTGATGACTCGTGTGTGTTGCTTAAGACTCTTCCACGCCCT 600  
 Db 370 ProGluProPheAlaLysLeuMetLysGluCysTrpGlnGlnAspProHisIleArgPro 389  
 QY 601 TCATGAGGAATTTGTGAATAATGACT 630  
 Db 390 SerPheAlaLeuIleLeuGluGlnLeuThr 399

RESULT 12

US-10-153-668-436  
 ; Sequence 436, Application US/10153668  
 ; Publication No. US20030092616A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HONDA, Goichi  
 ; APPLICANT: MATSUDA, Akio  
 ; APPLICANT: MURAMATSU, Shuji  
 ; APPLICANT: ISHIZAWA, Kenya  
 ; TITLE OF INVENTION: STAT6 Activating Gene  
 ; FILE REFERENCE: 1254-0207P  
 ; CURRENT APPLICATION NUMBER: US/10/153,668  
 ; CURRENT FILING DATE: 2002-05-24  
 ; PRIOR APPLICATION NUMBER: US 60/293,172  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: US 60/316,031  
 ; PRIOR FILING DATE: 2001-08-31  
 ; PRIOR APPLICATION NUMBER: US 60/328,403  
 ; PRIOR FILING DATE: 2001-10-12  
 ; PRIOR APPLICATION NUMBER: JP 2001-157043  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: JP 2001-260681  
 ; PRIOR FILING DATE: 2001-08-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-313175  
 ; NUMBER OF SEQ ID NOS: 488  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 436  
 ; LENGTH: 746  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-153-668-436

Alignment Scores:  
 Pred. No.: 7,41e-29 Length: 746  
 Score: 353.50 Matches: 75  
 Percent Similarity: 54.63% Conservative: 43  
 Best Local Similarity: 34.72% Mismatches: 81  
 Query Match: 28.23% Indels: 17

DB: 15 Gaps: 8  
 US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-153-668-436 (1-746)  
 QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGAGCGCTGC 63  
 Db 194 AspIleTyrHisLeuArgLysLeuLysHisProAsnIleIleThrPheLysGlyValCys 213  
 QY 64 TTGAATCCA-----GTGTGCTCTTCTGATGGAATAATGCTGAAGGGGCTCTTTATATAAT 117  
 Db 214 ThrGlnAlaProCysTyrCysIleLeuMetGluPheCysAlaGlnGlyGlnLeuTyrGlu 233  
 QY 118 GTGTGTCATGTGCTGAACCATTCATATATATATCTGCTGCCCGCCATGATGTGGTGT 177  
 Db 234 ValLeuArgAlaGlyArgProVal-----ThrProSerLeuLeuValAspTrpSer 250  
 QY 178 TTACAGTGTTCACAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCAC 237  
 Db 251 MetGlyIleAlaGlyGlyMetAsnTyrLeuHisLeuHisLys-----IleIleHis 267  
 QY 238 AGGGACCTGAAACCCAACTTACTGCTGTGTCAGGGGGGACAGTCTCTAAATAATTTGT 297  
 Db 268 ArgAspLeuLysSerProAsn---MetLeuIleThrTyrAspValValLysIleSer 286  
 QY 298 GATTTTGTACAGCCTGTGACATTCAGACACATGACCAATACAAAG-----GGAGT 351  
 Db 287 AspPheGlyThrSerLysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThr 306  
 QY 352 GCTGCTGTGATGGCAGCTCAAGCTTTTGAAGTAGTAATTCAGTGAATAATTCAGCTGC 411  
 Db 307 ValAlaTrpMetAlaProGluValIleArgAsnGluProValSerGluLysValAspIle 326  
 QY 412 TTCAGCTGGGTATTATCTTTGGGAAGTATAACCGCTCGGAACCCCTTTGATGAGATT 471  
 Db 327 TrpSerPheGlyValValLeuTrpGluLeuLeuThrGlyGluIleProTyrLysAspVal 346  
 QY 472 GTGGCCCGAGCTTCCGAATCATGTGGCTGT---CATATGCTGCTCGACCCACTG 528  
 Db 347 AspSerSerAla-----IleIleTrpGlyValGlySerAsnSerLeuHisLeuProVal 364  
 QY 529 ATAAAAATTTACCTAAGCCCATTCAGAGCTGATGCTGCTGTGTTGGTCTAAAGATCCT 588  
 Db 365 ProSerSerCysProAspGlyPheLysIleLeuLeuArgLysTrpAsnSerLysPro 384  
 QY 589 TCCCGAGCGCTTCAATGGAGGAATTTGTGAATAATATGACTCACTG 636  
 Db 385 ArgAsnArgProSerPheArgGln-----IleLeuLeuHisLeu 397

RESULT 13

US-10-153-668-324  
 ; Sequence 324, Application US/10153668  
 ; Publication No. US20030092616A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HONDA, Goichi  
 ; APPLICANT: MATSUDA, Akio  
 ; APPLICANT: MURAMATSU, Shuji  
 ; APPLICANT: ISHIZAWA, Kenya  
 ; TITLE OF INVENTION: STAT6 Activating Gene  
 ; FILE REFERENCE: 1254-0207P  
 ; CURRENT APPLICATION NUMBER: US/10/153,668  
 ; CURRENT FILING DATE: 2002-05-24  
 ; PRIOR APPLICATION NUMBER: US 60/293,172  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: US 60/316,031  
 ; PRIOR FILING DATE: 2001-08-31  
 ; PRIOR APPLICATION NUMBER: US 60/328,403  
 ; PRIOR FILING DATE: 2001-10-12  
 ; PRIOR APPLICATION NUMBER: JP 2001-157043  
 ; PRIOR FILING DATE: 2001-05-25  
 ; PRIOR APPLICATION NUMBER: JP 2001-260681  
 ; PRIOR FILING DATE: 2001-08-30  
 ; PRIOR APPLICATION NUMBER: JP 2001-313175  
 ; PRIOR FILING DATE: 2001-10-10



; NUMBER OF SEQ ID NOS: 488  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 324  
; LENGTH: 859  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-153-668-324

Alignment Scores:  
Pred. No.: 7,67e-29 Length: 859  
Score: 353.50 Matches: 75  
Percent Similarity: 54.63% Conservative: 43  
Best Local Similarity: 34.72% Mismatches: 81  
Query Match: 28.23% Indels: 17  
DB: 15 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-153-668-324 (1-859)

```
QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCCTGC 63
Db 161 AspilelyshisLeuArgLysLeuLysHisProAsnIleIleThrPheLysGlyValCys 180
QY 64 TTGAATCCA-----GTGTCCTGTGTGATGATGCTGAAGGGGCTCTTTATATAAT 117
Db 181 ThrGlnAlaProCysTyrCysIleLeuMetGluPheCysAlaGlnGlyGlnLeuTyrGlu 200
QY 118 GTGCTGCATGTGTGTAACCATTCCTATTAATCTGCTGCCACGCAATGAGTTGGTGT 177
Db 201 ValLeuArgAlaGlyArgProVal-----ThrProSerLeuLeuValAspTyrSer 217
QY 178 TTACAGTGTTCCTCAAGAGTGTGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCAC 237
Db 218 MetGlyIleAlaGlyGlyMetAsnTyrLeuHisLeuHisLys-----IleIleHis 234
QY 238 AGGACCTGAAACACCACTTACTGCTGTGTGCGGGGAGCAGTCTAAATAATTGT 297
Db 235 ArgAspLeuLysSerProAsn--MetLeuIleThrTyrAspValValLysIleSer 253
QY 298 GATTTTGGTACAGCTGTGACATTCACACACATGACCAATACAAAG-----GGAGT 351
Db 254 AspPheGlyThrSerLysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThr 273
QY 352 GCTGCTTGGATGGCACCCTGAAGTTTTCGAAGTAGTAATTACAGTGAATAATGTGAGTGC 411
Db 274 ValAlaTyrMetAlaProGluValIleArgAsnGluProValSerGluLysValAspIle 293
QY 412 TTACAGTGGGGTATTATCTTTGGGAAGTGAACGCGTGGAAACCCCTTTGATGAGATT 471
Db 294 TrpSerPheGlyValValLeuTyrGluLeuThrGlyGluIleProTyrLysAspVal 313
QY 472 GGTGGCCAGCTTCCGAATCATGTGGCTGTT---CATATGGTACTCGACCACTG 528
Db 314 AspSerSerAla-----IleIleTyrGlyValGlySerAsnSerLeuHisLeuProVal 331
QY 529 ATAAAAAATTACCTAAGCCCATTCAGAGCTGATGACTCGTTTGGTCTAAAGATCCT 588
Db 332 ProSerSerCysProAspGlyPheLysIleLeuLeuArgGlnCysTrpAsnSerLysPro 351
QY 589 TCCACGGCCCTTCATGGAGAAATGTGAAATAATGACTCCTTG 636
Db 352 ArgAsnArgProSerPheArgGln-----IleLeuLeuHisLeu 364
```

RESULT 14

US-10-153-668-438  
; Sequence 438, Application US/10153668  
; Publication No. US20030092616A1  
; GENERAL INFORMATION:  
; APPLICANT: HONDA, Goichi  
; APPLICANT: MATSUDA, Akio  
; APPLICANT: MURAMATSU, Shuji  
; APPLICANT: ISHIZAWA, Kenya  
; TITLE OF INVENTION: STAT6 Activating Gene  
; FILE REFERENCE: 1254-0207P

; CURRENT APPLICATION NUMBER: US/10/153,668  
; CURRENT FILING DATE: 2002-05-24  
; PRIOR APPLICATION NUMBER: US 60/293,172  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: US 60/316,031  
; PRIOR FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: US 60/328,403  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: JP 2001-157043  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: JP 2001-260681  
; PRIOR FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: JP 2001-313175  
; PRIOR FILING DATE: 2001-10-10  
; NUMBER OF SEQ ID NOS: 488  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 438  
; LENGTH: 892  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-153-668-438

Alignment Scores:  
Pred. No.: 7,74e-29 Length: 892  
Score: 353.50 Matches: 75  
Percent Similarity: 54.63% Conservative: 43  
Best Local Similarity: 34.72% Mismatches: 81  
Query Match: 28.23% Indels: 17  
DB: 15 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-153-668-438 (1-892)

```
QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCCTGC 63
Db 194 AspilelyshisLeuArgLysLeuLysHisProAsnIleIleThrPheLysGlyValCys 213
QY 64 TTGAATCCA-----GTGTCCTGTGTGATGATGCTGAAGGGGCTCTTTATATAAT 117
Db 214 ThrGlnAlaProCysTyrCysIleLeuMetGluPheCysAlaGlnGlyGlnLeuTyrGlu 233
QY 118 GTGCTGCATGTGTGTAACCATTCCTATTAATCTGCTGCCACGCAATGAGTTGGTGT 177
Db 234 ValLeuArgAlaGlyArgProVal-----ThrProSerLeuLeuValAspTyrSer 250
QY 178 TTACAGTGTTCCTCAAGAGTGTGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCAC 237
Db 251 MetGlyIleAlaGlyGlyMetAsnTyrLeuHisLeuHisLys-----IleIleHis 267
QY 238 AGGACCTGAAACACCACTTACTGCTGTGTGCGGGGAGCAGTCTTAAATAATTGT 297
Db 268 ArgAspLeuLysSerProAsn--MetLeuIleThrTyrAspValValLysIleSer 286
QY 298 GATTTTGGTACAGCTGTGACATTCACACACATGACCAATAACAAG-----GGAGT 351
Db 287 AspPheGlyThrSerLysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThr 306
QY 352 GCTGCTTGGATGGCACCCTGAAGTTTTCGAAGTAGTAATTACAGTGAATAATGTGAGTGC 411
Db 307 ValAlaTyrMetAlaProGluValIleArgAsnGluProValSerGluLysValAspIle 326
QY 412 TTACAGTGGGGTATTATCTTTGGGAAGTGAACCGCTGCGGAAACCCCTTTGATGAGATT 471
Db 327 TrpSerPheGlyValValLeuTyrGluLeuThrGlyGluIleProTyrLysAspVal 346
QY 472 GGTGGCCAGCTTCCGAATCATGTGGCTGTT---CATATGGTACTCGACCACTG 528
Db 347 AspSerSerAla-----IleIleTyrGlyValGlySerAsnSerLeuHisLeuProVal 364
QY 529 ATAAAAAATTACCTAAGCCCATTCAGAGCTGATGACTCGTTTGGTCTAAAGATCCT 588
Db 365 ProSerSerCysProAspGlyPheLysIleLeuLeuArgGlnCysTrpAsnSerLysPro 384
QY 589 TCCACGGCCCTTCATGGAGAAATGTGAAATAATGACTCCTTG 636
```

```
Db 385 ArgAsnArgProSerPheArgGln.....lleLeuLeuHisLeu 397
|||||
231 ValLeuSerGlyLysArgIleProProAspIle-----LeuValAsnTrpAla 246
|||||
178 TTACAGTGTCCCAAGAGTGGCTTATCTTCACAGCATGCCAACCGCGTAATTAC 237
|||||
247 ValGlnIleAlaArgGlyMetAsnTrpLeuLeuAspGluAlaIleValProIleHis 266
|||||
238 AGGGACCTGAACACCAACTTACTGCTGTT-----GCAGG 276
|||||
267 ArgAspLeuLysSerSerAsnIleLeuIleLeuGlnLysValGluAsnGlyAspLeuSer 286
|||||
277 GGCAGAGTTCTAAATAATTTGCAITTTGGTACAGCTGTGAC---ATTGACACACATG 333
|||||
287 AsnLysIleLeuLysIleThrAspPheGlyLeuAlaArgGluTrpHisArgThrLys 306
|||||
334 ACCAATAACACAGGGAGTCTGCTGGATGGCACCTGAAGTTTTCGAAGGTAGTATTAC 393
|||||
307 MetSerAlaAlaGlyThrTrpAlaTrpMetAlaProGluValIleArgAlaSerMetPhe 326
|||||
394 AGTGAATAATGTGACCTCTTCAGCTGGGGTATTATTCTTTGGGAGTGATAAGCGCTCG 453
|||||
327 SerLysGlySerAspValTrpSerTrpGlyValLeuLeuTrpGluLeuLeuThrGlyGlu 346
|||||
454 AAACCTTTGATGAGATTGGTGGCCAGCTTCCGAGTATCATGGGCTGTTCTATATGTT 513
|||||
347 ValProPheArgGlyIleAspGlyLeuAlaValAlaTrpGlyValAlaMet---AsnLys 365
|||||
514 ACTGACACCACTGATATAAAATTTACCTAAGCCCATTTGAGAGCCTGATGACTCGTTGT 573
|||||
366 LeuAlaLeuProIleProSerThrCysProGluProPheAlaLysLeuMetGluAspCys 385
|||||
574 TGCTCTAAAGATCTCTCCAGCGCCCTTCAATGGAGGAAATTTGAAAAATAATGACT 630
|||||
386 TrpAsnProAspProHisSerArgProSerPheThrAsnIleLeuAspGlnLeuThr 404
|||||
```

Search completed: December 4, 2003, 09:28:21  
Job time : 127.703 secs

```
Db 385 ArgAsnArgProSerPheArgGln.....lleLeuLeuHisLeu 397
|||||
191 GluAlaLysLeuPheAlaMetLeuLysHisProAsnIleLeuAlaLeuArgGlyValCys 210
|||||
64 TTGAATCCA-----GTGTGCTTGTGATGAATATGCTGAAGGGCTCTTTATATAAT 117
|||||
211 LeuLysGluProAsnLeuCysLeuValMetGluPheAlaArgGlyGlyProLeuAsnArg 230
|||||
118 GTGCTGCTGCTGTGCTGAACCATTCGCATATTATATCTGCTGCCACGCAATGAGTTGTTGT 177

RESULT 15
US-10-288-798-12
; Sequence 12, Application US/10288798
; Publication No. US20030207299A1
; GENERAL INFORMATION:
; APPLICANT: BANDMAN, Olga; NGUYEN, Damiel B.;
; APPLICANT: WALIA, Narinder K.; HAFALIA, April J.A.;
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;
; APPLICANT: PATTERSON, Chandra; YUE, Henry;
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;
; APPLICANT: LU, Yan; ISON, Craig H.;
; APPLICANT: AZIMZAI, Yalda; BURRILL, John D.;
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;
; APPLICANT: LU, Dying Alina M.; LAL, Preeti G.;
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0209 USA
; CURRENT APPLICATION NUMBER: US/10/288,798
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 1097
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030207299A1 55052990CD1
US-10-288-798-12

Alignment Scores:
Pred. No.: 2.18e-28 Length: 1097
Score: 349.50 Matches: 82
Percent Similarity: 53.42% Conservative: 35
Best Local Similarity: 37.44% Mismatches: 87
Query Match: 27.92% Indels: 15
DB: 12 Gaps: 5

US-09-830-144-1_COPY_408_1091 (1-684) x US-10-288-798-12 (1-1097)
QY 4 GAGCTTCGGCAGTTATCCGCTGTGACCATCTTAATTTGTAAGCTTTATGGAGCTGC 63
Db 191 GluAlaLysLeuPheAlaMetLeuLysHisProAsnIleLeuAlaLeuArgGlyValCys 210
QY 64 TTGAATCCA-----GTGTGCTTGTGATGAATATGCTGAAGGGCTCTTTATATAAT 117
Db 211 LeuLysGluProAsnLeuCysLeuValMetGluPheAlaArgGlyGlyProLeuAsnArg 230
QY 118 GTGCTGCTGCTGTGCTGAACCATTCGCATATTATATCTGCTGCCACGCAATGAGTTGTTGT 177
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: December 4, 2003, 09:00:20 ; Search time 20.7973 Seconds  
(without alignments)  
2783.114 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091  
Perfect score: 1252  
Sequence: 1 gttagagcttcggcagttatc.....cattacagatcctgtgtcag 684

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues  
Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framed\_n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09830144/runat\_03122003\_122345\_21337/app.query.fasta\_1.1230  
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPECL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptb -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 @CGN\_1\_1\_29 @runat\_03122003\_122345\_21337 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	579	4	US-09-529-279-4
2	1252	100.0	579	4	US-10-158-895-4
3	1252	100.0	590	4	US-09-529-279-15
4	1252	100.0	590	4	US-10-158-895-15
5	371	29.6	455	3	US-09-221-235-5
6	371	29.6	455	3	US-09-221-928-5
7	371	29.6	455	3	US-09-221-928-5
8	371	29.6	455	3	US-09-221-236-5
9	371	29.6	455	3	US-09-221-416-5
10	371	29.6	455	3	US-09-221-245-5
11	371	29.6	455	3	US-09-163-115-5
12	371	29.6	455	3	US-09-221-528-5

13	371	29.6	455	3	US-09-593-553-5	Sequence 5, Appli
14	371	29.6	455	3	US-09-221-237-5	Sequence 5, Appli
15	371	29.6	455	3	US-09-399-588-2	Sequence 2, Appli
16	363.5	29.0	394	4	US-09-345-473E-19	Sequence 19, Appl
17	353.5	28.2	668	1	US-08-205-018-2	Sequence 2, Appli
18	353.5	28.2	859	1	US-08-395-580-2	Sequence 2, Appli
19	353.5	28.2	859	5	PCT-US95-02792-2	Sequence 2, Appli
20	346	27.6	328	4	US-09-345-473E-18	Sequence 18, Appli
21	338.5	27.0	835	3	US-09-291-839-2	Sequence 2, Appli
22	338.5	27.0	835	4	US-09-458-457-2	Sequence 2, Appli
23	335.5	26.8	835	4	US-09-458-457-2	Sequence 8, Appli
24	323	25.8	821	1	US-07-928-464-2	Sequence 2, Appli
25	323	25.8	821	1	US-08-003-311B-2	Sequence 2, Appli
26	323	25.8	821	1	US-08-261-432-2	Sequence 2, Appli
27	323	25.8	821	5	PCT-US93-07347-2	Sequence 2, Appli
28	320	25.6	263	3	US-09-035-706-5	Sequence 5, Appli
29	320	25.6	263	3	US-08-955-841-5	Sequence 5, Appli
30	320	25.6	263	4	US-09-390-425-5	Sequence 5, Appli
31	320	25.6	263	4	US-09-566-906-5	Sequence 5, Appli
32	309.5	24.7	269	2	US-07-857-224B-79	Sequence 79, Appl
33	307.5	24.6	276	2	US-07-857-224B-72	Sequence 72, Appl
34	306	24.4	275	2	US-07-857-224B-71	Sequence 71, Appl
35	305.5	24.4	786	4	US-09-509-802-2	Sequence 409, App
36	305.5	24.4	786	4	US-09-312-283C-409	Sequence 2, Appli
37	305.5	24.4	787	3	US-09-188-930-334	Sequence 334, App
38	305.5	24.4	787	4	US-09-312-283C-334	Sequence 334, App
39	304.5	24.3	304	2	US-08-701-191A-27	Sequence 27, Appl
40	302.5	24.2	316	1	US-08-278-089A-16	Sequence 16, Appl
41	302.5	24.2	316	2	US-08-388-957A-15	Sequence 15, Appl
42	299.5	23.9	527	4	US-08-426-509A-10	Sequence 10, Appl
43	299.5	23.9	527	4	US-08-232-545-10	Sequence 10, Appl
44	299.5	23.9	527	5	PCT-US95-05008-10	Sequence 10, Appl
45	299.5	23.9	625	1	US-08-391-615-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-09-529-279-4  
; Sequence 4, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-4

Alignment Scores:  
Pred. No.: 4.53e-145 Length: 579  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-529-279-4 (1-579)

QY 1 GTAGAGCTTCGGCAGTTATCCCTGTGAACCATCTTAATTTGTAAGCTTTATGAGCC 60

Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyAla 95  
QY 61 TCGTGAATCCAGTGTGCTGTGATGGAATATGCTGAAGGGGCTCTTATATATATG 120  
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115  
QY 121 CTGCTGGTCTCAACCATGTCATATATCTGCTGCCACCAATGATGCTGCTTTA 180  
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
QY 181 CAGTGTTCACAGAGTGGCTTATCTTCACAGATGCAACCAAGCGCTAAATTCACAGG 240  
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg 155  
QY 241 GACCTGAACCAACCAACTTACTGCTGGTGGAGGGGAGCAGTCTTAAATAATTTGAT 300  
Db 156 AspLeuLysProAsnLeuLeuValAlaGlyThrValLeuLysLeuCysAsp 175  
QY 301 TTTGGTACAGCTGTGACATTCACACATGACCAATTAACAGGGAGTGGCTTGG 360  
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195  
QY 361 ATGGCCACTGAAGTTTTGAAGTAGTAATTAATGAGTGAATAATGTCAGCTCTTCAGCTGG 420  
Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
QY 421 GGTATTATCTTTGGGAAGTGATAACCGCTCGGAACCCCTTGTGATGAGTGGGCCCA 480  
Db 216 GlyIleLeuTrpGluValIleThrArgLysProPheAspGluIleGlyPro 235  
QY 481 GCCTTCGGAATCATGGCTGTGCTAATGCTGACCTGACCAACCACTGATAAAAAATTTA 540  
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProLeuLysLeuLysAsnLeu 255  
QY 541 CCTAAGCCCATTCAGACCTGATGACCTGCTGTTGTTGTTAAAGATCTTCCAGCGCCCT 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGAGGAATGTGAAATAATGACTCATCTGATGCGGTACTTCCAGAGCGAGAT 660  
Db 276 SerMetGluGluLeuValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
QY 661 GAGCCATTACAGTATCCTTGTCTAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303

## RESULT 2

US-10-158-895-4  
; Sequence 4, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; PRIOR FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-895-4  
Alignment Scores:

Pred. No.: 4,53e-145 Length: 579  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-158-895-4 (1-579)  
QY 1 GTAGACTTCGGCAGTATCCCGTGTGAACCATCTTAATATCTAAAGCTTTATGAGCC 60  
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyAla 95  
QY 61 TCGTGAATCCAGTGTGCTGTGATGGAATATGCTGAAGGGGCTCTTATATATATG 120  
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115  
QY 121 CTGCTGGTCTCAACCATGTCATATATCTGCTGCCACCAATGATGCTGCTTTA 180  
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
QY 181 CAGTGTTCACAGAGTGGCTTATCTTCACAGATGCAACCAAGCGCTAAATTCACAGG 240  
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg 155  
QY 241 GACCTGAACCAACCAACTTACTGCTGGTGGAGGGGAGCAGTCTTAAATAATTTGAT 300  
Db 156 AspLeuLysProAsnLeuLeuValAlaGlyThrValLeuLysLeuCysAsp 175  
QY 301 TTTGGTACAGCTGTGACATTCACACATGACCAATTAACAGGGAGTGGCTTGG 360  
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195  
QY 361 ATGGCCACTGAAGTTTTGAAGTAGTAATTAATGAGTGAATAATGTCAGCTCTTCAGCTGG 420  
Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
QY 421 GGTATTATCTTTGGGAAGTGATAACCGCTCGGAACCCCTTGTGATGAGTGGGCCCA 480  
Db 216 GlyIleLeuTrpGluValIleThrArgLysProPheAspGluIleGlyPro 235  
QY 481 GCCTTCGGAATCATGGCTGTGCTAATGCTGACCTGACCAACCACTGATAAAAAATTTA 540  
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProLeuLysLeuLysAsnLeu 255  
QY 541 CCTAAGCCCATTCAGACCTGATGACCTGCTGTTGTTGTTAAAGATCTTCCAGCGCCCT 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGAGGAATGTGAAATAATGACTCATCTGATGCGGTACTTCCAGAGCGAGAT 660  
Db 276 SerMetGluGluLeuValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
QY 661 GAGCCATTACAGTATCCTTGTCTAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303

## RESULT 3

US-09-529-279-15  
; Sequence 15, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22

; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-15

Alignment Scores:  
Pred. No.: 4.57e-145 Length: 590  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-529-279-15 (1-590)

```
QY 1 GTAGAGCTTCGGCAGTTATCCGCTGTGAACCATCTCTAATATTGTAAAGCTTTATGAGCC 60
Db ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95
QY 61 TCGTTGAATCCAGTGTGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 120
Db CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115
QY 121 CTGCATGCTGTGAACCATTCGCATATATATCTGCTGCCAGGCATGAGTTGGTGTTA 180
Db LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135
QY 181 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCCGCTAATTCACAGG 240
Db GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155
QY 241 GACCTGAAACCAACCAACTTACTGCTGTTCAGGGGGAGAGTTCTAAAAATTTGTGAT 300
Db AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175
QY 301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATACAGGGAGTGTGCTTGG 360
Db PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195
QY 361 ATGGCCTCAAGTTTTTGAAGGTAGTAATACAGTGAATAATGTAGCTCTTCAGCTGG 420
Db 196 MetAlaProGluValPheGluGlySerAsnLysSerGlnLysCysAspValPheSerTrp 215
QY 421 GGTATTATCTTTGGGAGTGAATAACGCTCGGAAACCTTTGATGAGATTGGTGGCCCA 480
Db 216 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 235
QY 481 GCTTTCCGAATCATGTGGCTGTTCATATGTTACTCGACACACACTGATAAAAATTTA 540
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255
QY 541 CCTAAGCCCATTCAGAGCTGATGACTCGTTGGTCTAAGATCCCTCCAGCGCCCT 600
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275
QY 601 TCAATGAGGAAATTTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 660
Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295
QY 661 GAGCCATTACAGTATCTTGTTCAG 684
Db 296 GluProLeuGlnTyrProCysGln 303
```

RESULT 4

; US-10-158-895-15  
; Sequence 15, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO

; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-895-15

Alignment Scores:  
Pred. No.: 4.57e-145 Length: 590  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-10-158-895-15 (1-590)

```
QY 1 GTAGAGCTTCGGCAGTTATCCGCTGTGAACCATCTAATATTGTAAAGCTTTATGAGCC 60
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95
QY 61 TCGTTGAATCCAGTGTGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 120
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115
QY 121 CTGCATGCTGTGAACCATTCGCATATATATCTGCTGCCAGGCATGAGTTGGTGTTA 180
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135
QY 181 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCCGCTAATTCACAGG 240
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155
QY 241 GACCTGAAACCAACCAACTTACTGCTGTTCAGGGGGAGAGTTCTAAAAATTTGTGAT 300
Db 156 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175
QY 301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATACAGGGAGTGTGCTTGG 360
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195
QY 361 ATGGCCTCAAGTTTTTGAAGGTAGTAATACAGTGAATAATGTAGCTCTTCAGCTGG 420
Db 196 MetAlaProGluValPheGluGlySerAsnLysSerGlnLysCysAspValPheSerTrp 215
QY 421 GGTATTATCTTTGGGAGTGAATAACGCTCGGAAACCTTTGATGAGATTGGTGGCCCA 480
Db 216 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 235
QY 481 GCTTTCCGAATCATGTGGCTGTTCATATGTTACTCGACACACACTGATAAAAATTTA 540
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255
QY 541 CCTAAGCCCATTCAGAGCTGATGACTCGTTGGTCTAAGATCCCTCCAGCGCCCT 600
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275
QY 601 TCAATGAGGAAATTTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 660
Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295
```

```
QY 661 GAGCATTACAGTACCTTGTCTGAG 684
Db 296 GluProLeuGlnTyrProCysGln 303

RESULT 5
US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-235-5

Alignment Scores:
Pred. No.: 9,61e-37 Length: 455
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 3 Gaps: 8

US-09-830-144-1_COPY_408_1091 (1-684) x US-09-221-235-5 (1-455)
QY 4 GAGCTTCGGCAGTATCCCGTGTGACCATCTCTAAATTTGTAAGCTTTATGGAGCCTGC 63
Db 53 GluAlaGluLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 72
QY 64 TTGAATCCA-----GTGTCCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db 73 LeuGluProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92
QY 118 GTGCTGCATGGT-----GCTGAACCATTCGCATATTATATCTGCTGCCCGCAATGAGT 171
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109
QY 172 TGGTGTTTACAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTA 231
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129
QY 232 ATTCACAGGACCTGAAACCAACCAACTTACTGCTGTTGTCAGGGGGGACAGTTCTAAAA 291
Db 130 IleHisArgAspLeuLysSerArgAsnValIleAlaAlaAspGly---ValLeuLys 148
QY 292 ATTGTGATTTTGGTACAGCTGTGACATTCAGACACACATGACCAATAACAAG----- 345
Db 149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167
QY 346 GGGAGTCTCTCTGGATGGCAGCTGAAAGTTTTTGAAGGTAGTAGTAATTACAGTGAATAATGT 405
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187
QY 406 GAGCTCTTACAGTGGGTATATCTTTGGAGAGTGAATACCGGTGCGAAACCCCTTTGAT 465
Db 188 AspThrTyrSerTyrGlyValValLeuTrpGluMetLeuThrArgGluValProPheLys 207
QY 466 GAGATTGGTGGCCAGCTTTCCGAATCATGTGG---GCTGTTCAATAATGGTACTCTGACCA 522
Db 208 GlyLeuGluGly-----LeuGlnValAlaTrpLeuValValGluLysAsnGluArgLeu 225
QY 523 CCACGTGATAAAAAATTTACCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGTCTPAAA 582
Db 226 ThrIleProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245

US-09-830-144-1_COPY_408_1091 (1-684) x US-09-221-928-5 (1-455)
QY 583 GATCCTTCCAGCGCCCTTCAATGGAGAAATTTGAAAAATATG 627
Db 246 AspAlaLysLysArgProSerPheLysGlnIleIleSerIleLeu 260

RESULT 6
US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-928-5

Alignment Scores:
Pred. No.: 9,61e-37 Length: 455
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 3 Gaps: 8

US-09-830-144-1_COPY_408_1091 (1-684) x US-09-221-928-5 (1-455)
QY 4 GAGCTTCGGCAGTATCCCGTGTGACCATCTCTAAATTTGTAAGCTTTATGGAGCCTGC 63
Db 53 GluAlaGluLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 72
QY 64 TTGAATCCA-----GTGTCCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db 73 LeuGluProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92
QY 118 GTGCTGCATGGT-----GCTGAACCATTCGCATATTATATCTGCTGCCCGCAATGAGT 171
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109
QY 172 TGGTGTTTACAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTA 231
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129
QY 232 ATTCACAGGACCTGAAACCAACCAACTTACTGCTGTTGTCAGGGGGGACAGTTCTAAAA 291
Db 130 IleHisArgAspLeuLysSerArgAsnValIleAlaAlaAspGly---ValLeuLys 148
QY 292 ATTGTGATTTTGGTACAGCTGTGACATTCAGACACACATGACCAATAACAAG----- 345
Db 149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167
QY 346 GGGAGTCTCTCTGGATGGCAGCTGAAAGTTTTTGAAGGTAGTAGTAATTACAGTGAATAATGT 405
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187
QY 406 GAGCTCTTACAGTGGGTATATCTTTGGAGAGTGAATACCGGTGCGAAACCCCTTTGAT 465
Db 188 AspThrTyrSerTyrGlyValValLeuTrpGluMetLeuThrArgGluValProPheLys 207
QY 466 GAGATTGGTGGCCAGCTTTCCGAATCATGTGG---GCTGTTCAATAATGGTACTCTGACCA 522
Db 208 GlyLeuGluGly-----LeuGlnValAlaTrpLeuValValGluLysAsnGluArgLeu 225
QY 523 CCACGTGATAAAAAATTTACCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGTCTPAAA 582
Db 226 ThrIleProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245
```

```
Db      226 ThrileProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245
Qy      583 GATCCTTCCAGCCCTTCAATGAGGAAATGTTGAAAAATATG 627
Db      246 AspalalysLysArgProSerPheLysGlnIleSerIleLeu 260
RESULT 7
US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER FILING DATE:
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Alignment Scores:
Pred. No.:      9,61e-37      Length:      455
Score:          371.00      Matches:      81
Percent Similarity: 56.28%      Conservative: 40
Best Local Similarity: 37.67%      Mismatches: 80
Query Match:     29.63%      Indels:      14
DB:              3          Gaps:         8

US-09-830-144-1_COPY_408_1091 (1-684) x US-09-221-527-5 (1-455)
Qy      4 GAGCTTCGGCAGTATCCCGTGTGAACCATCTAATATTGTAAAGCTTTATGGAGCCTGC 63
Db      53 GlualaGluIleLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 72
Qy      64 TTGAATCCA-----GTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db      73 LeuGluProProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92
Qy      118 GTGCTGATGCT-----GCTGAACCATTCGCATTTATATCTGCTGCCCGCAATGAGT 171
Db      93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109
Qy      172 TGGTGTGTTACAGTGTCCAGGAGTGGCTTATCTTACAGCATGCAACCCAAAGCGCTA 231
Db      110 TrpAlaThrAspValAlaLysGlyMetHisTyrIleHisMetGluAlaProValIysVal 129
Qy      232 ATTCACAGGGACCTGAAACCAACCAACTTACTGCTGTGTCAGGGGGGACAGTTCATAAA 291
Db      130 IleHisArgAspLeuLysSerArgAsnValIleAlaAlaAspGly---ValLeuLys 148
Qy      292 ATTTGTGATTTTGTACAGCTGTGACATTCAGACACATGACCAATAACAAG----- 345
Db      149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167
Qy      346 GGGAGTGTGCTGGATGGCACCTGAAGTCTTTTGAAGTAGTAGTAATACAGTGAATAATGT 405
Db      168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187
Qy      406 GACGCTCTTACGTGGGTATTATCTTTGGGAAGTGAACCGCTCGGAAACCCCTTTGAT 465
Db      188 AspThrTyrSerTyrGlyValValIleuTrpGluMetLeuThrArgGluValProPheLys 207
Qy      466 GAGATTGTGTCGCCAGCTTTCCGAATCATGTGG---GCTGTTCTATAATGGTACTCGACCA 522
Db      208 GlyLeuGluGly-----LeuGlnValAlaTrpLeuValValGluLysAsnGluArgLeu 225
Qy      523 CCACTGATAAAAAATTTTACCTAAGCCCATTCGAGAGCCTGATGACTCGTTGTTGCTCTAAA 582
```

```
Db      226 ThrileProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245
Qy      583 GATCCTTCCAGCCCTTCAATGAGGAAATGTTGAAAAATATG 627
Db      246 AspalalysLysArgProSerPheLysGlnIleSerIleLeu 260
RESULT 8
US-09-221-236-5
; Sequence 5, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER FILING DATE: 1998-12-28
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-236-5

Alignment Scores:
Pred. No.:      9,61e-37      Length:      455
Score:          371.00      Matches:      81
Percent Similarity: 56.28%      Conservative: 40
Best Local Similarity: 37.67%      Mismatches: 80
Query Match:     29.63%      Indels:      14
DB:              3          Gaps:         8

US-09-830-144-1_COPY_408_1091 (1-684) x US-09-221-236-5 (1-455)
Qy      4 GAGCTTCGGCAGTATCCCGTGTGAACCATCTAATATTGTAAAGCTTTATGGAGCCTGC 63
Db      53 GlualaGluIleLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 72
Qy      64 TTGAATCCA-----GTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db      73 LeuGluProProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92
Qy      118 GTGCTGATGCT-----GCTGAACCATTCGCATTTATATCTGCTGCCCGCAATGAGT 171
Db      93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109
Qy      172 TGGTGTGTTACAGTGTCCAGGAGTGGCTTATCTTACAGCATGCAACCCAAAGCGCTA 231
Db      110 TrpAlaThrAspValAlaLysGlyMetHisTyrIleHisMetGluAlaProValIysVal 129
Qy      232 ATTCACAGGGACCTGAAACCAACCAACTTACTGCTGTGTCAGGGGGGACAGTTCATAAA 291
Db      130 IleHisArgAspLeuLysSerArgAsnValIleAlaAlaAspGly---ValLeuLys 148
Qy      292 ATTTGTGATTTTGTACAGCTGTGACATTCAGACACATGACCAATAACAAG----- 345
Db      149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167
Qy      346 GGGAGTGTGCTGGATGGCACCTGAAGTCTTTTGAAGTAGTAGTAATACAGTGAATAATGT 405
Db      168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187
Qy      406 GACGCTCTTACGTGGGTATTATCTTTGGGAAGTGAACCGCTCGGAAACCCCTTTGAT 465
Db      188 AspThrTyrSerTyrGlyValValIleuTrpGluMetLeuThrArgGluValProPheLys 207
Qy      466 GAGATTGTGTCGCCAGCTTTCCGAATCATGTGG---GCTGTTCTATAATGGTACTCGACCA 522
Db      208 GlyLeuGluGly-----LeuGlnValAlaTrpLeuValValGluLysAsnGluArgLeu 225
```

QY 523 CCACCTGATAAAAAATTTACCTAAGCCATTGAGAGCCTGATGACTCGTCTGTCTCTAAA 582  
Db 226 ThrileProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245  
QY 583 GATCCTTCCAGCGCCCTTCAATGGAGGAATTTGTGAAAAATAATG 627  
Db 246 AspAlaLysLysArgProSerPheLysGlnIleSerIleLeu 260

## RESULT 9

US-09-221-416-5  
; Sequence 5, Application US/09221416  
; Patent No. 6153417  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221.416  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-416-5

Alignment Scores:  
Pred. No.: 9,61e-37 Length: 455  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 3 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-221-416-5 (1-455)

QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCTGC 63  
Db 53 GluAlaGluLeuLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValle 72  
QY 64 TTGAATCCA-----GTGTGCTTGTGATGAATGCTGAAGGGGCTCTTTATAAAT 117  
Db 73 LeuGluProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92  
QY 118 GTGCTGCATGCT-----GCTGAACCATTCGCATATTATATCTGCTGCCCGCAATGAGT 171  
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109  
QY 172 TGGTGTTCACGTGTTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTA 231  
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129  
QY 232 ATTCACAGGACCTGAAACCAACCACTTACTGCTGGTTCAGGGGGCAGACAGTTCTAAAA 291  
Db 130 IleHisArgAspLeuLysSerArgAsnValIleAlaAspGly---ValleuLys 148  
QY 292 ATTTGTGATTTTGTACAGCTGTGACATTCAGACACACATGACCAATAACAAG----- 345  
Db 149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrHisMetSerLeuVal 167  
QY 346 GGGAGTGTCTGCTGATGGCAGCTGAAGTTTTCAGGTAGTAGTAATTACAGTGAATAATGT 405  
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187  
QY 406 GACCTCTTCAGCTGGGGTATTATTCTTTGGGAAGTGAACGCTCGGAAACCCCTTTGAT 465  
Db 188 AspThrTyrSerTyrGlyValValLeuTrpGluMetLeuThrArgGluValProPheLys 207  
QY 466 GAGATTGTGGCCAGCTTTCGGAATCATGTGG---GCTGTTTCATATGTTACTCGACCA 522  
Db 208 GlyLeuGluGly-----LeuGlnValAlaTrpIleuValGluLysAsnGluArgLeu 225

QY 523 CCACCTGATAAAAAATTTACCTAAGCCATTGAGAGCCTGATGACTCGTCTGTCTAAA 582  
Db 226 ThrileProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245  
QY 583 GATCCTTCCAGCGCCCTTCAATGGAGGAATTTGTGAAAAATAATG 627  
Db 246 AspAlaLysLysArgProSerPheLysGlnIleSerIleLeu 260

## RESULT 10

US-09-221-245-5  
; Sequence 5, Application US/09221245  
; Patent No. 6180358  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221.245  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: US 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-245-5

Alignment Scores:  
Pred. No.: 9,61e-37 Length: 455  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 3 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-221-245-5 (1-455)

QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCTGC 63  
Db 53 GluAlaGluLeuLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValle 72  
QY 64 TTGAATCCA-----GTGTGCTTGTGATGAATGCTGAAGGGGCTCTTTATAAAT 117  
Db 73 LeuGluProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92  
QY 118 GTGCTGCATGCT-----GCTGAACCATTCGCATATTATATCTGCTGCCCGCAATGAGT 171  
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109  
QY 172 TGGTGTTCACGTGTTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTA 231  
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129  
QY 232 ATTCACAGGACCTGAAACCAACCACTTACTGCTGGTTCAGGGGGCAGACAGTTCTAAAA 291  
Db 130 IleHisArgAspLeuLysSerArgAsnValIleAlaAspGly---ValleuLys 148  
QY 292 ATTTGTGATTTTGTACAGCTGTGACATTCAGACACACATGACCAATAACAAG----- 345  
Db 149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrHisMetSerLeuVal 167  
QY 346 GGGAGTGTCTGCTGATGGCAGCTGAAGTTTTCAGGTAGTAGTAATTACAGTGAATAATGT 405  
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187  
QY 406 GACCTCTTCAGCTGGGGTATTATTCTTTGGGAAGTGAACGCTCGGAAACCCCTTTGAT 465  
Db 188 AspThrTyrSerTyrGlyValValLeuTrpGluMetLeuThrArgGluValProPheLys 207  
QY 466 GAGATTGTGGCCAGCTTTCGGAATCATGTGG---GCTGTTTCATATGTTACTCGACCA 522  
Db 208 GlyLeuGluGly-----LeuGlnValAlaTrpIleuValGluLysAsnGluArgLeu 225



Db 208 GlyLeuGluGly-----LeuGlnValAlaTrpLeuValValGluLysAsnGluArgLeu 225  
Qy 523 CCACGTGATAAAATTTACCTAACCCATTGAGAGCCTGATGACTCGTTGTTGCTCTAAA 582  
Db 226 ThrIleProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245  
Qy 583 GATCCTTCCAGCGCCCTTCAATGGAGGAAATTTGAAAAATAATG 627  
Db 246 AspAlaLysLysArgProSerPheLysGlnIleIleSerIleLeu 260

## RESULT 11

US-09-163-115-5  
; Sequence 5, Application US/09163115A  
; Patent No. 6183962  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/163,115A  
; CURRENT FILING DATE: 1998-09-29  
; EARLIER APPLICATION NUMBER: 09/163,115  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-163-115-5

Alignment Scores:  
Pred. No.: 9.61e-37 Length: 455  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 3 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-163-115-5 (1-455)

Qy 4 GAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCCTGC 63  
Db 53 GluAlaGluIleLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 72  
Qy 64 TTGAATCCA-----GTGTGCTTGTGATGGAATATGCTGAAGGGCTCTTTATATAAT 117  
Db 73 LeuGlnProProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92  
Qy 118 GTGCTGCATGCT-----GCTGAACCATTTGCCATATTATATCTGCTGCCCGCAATGAGT 171  
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109  
Qy 172 TGGTGTTCACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTA 231  
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129  
Qy 232 ATTCACAGGAGCCTGAAACCAACCAACTTACTGCTGTTGTCAGGGGGCAGAGTTCTAAA 291  
Db 130 IleHisArgAspLeuLysSerArgAsnValValIleAlaAspGly---ValLeuLys 148  
Qy 292 ATTTGTGATTTTGTACAGCCTGTGCATTCACACACACATGACCAATAACAAG----- 345  
Db 149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167  
Qy 346 GGGAGTGTCTGGATGGACCTGAAGTTTTTGAAGGTAGTAAATACAGTGAATAATGT 405  
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187  
Qy 406 GACGCTTCACGTGGGTATTATCTTTGGAGAGTAAACGCTCGGAAACCCCTTTGAT 465  
Db 188 AspThrTyrSerTyrGlyValValLeuTrpGluMetLeuThrArgGluValProPheLys 207  
Qy 466 GAGATTGTGGCCAGCTTTCCGAAATCATGTGG---GCTGTTTCATATGATGGTACTCGACCA 522  
Db 208 GlyLeuGluGly-----LeuGlnValAlaTrpLeuValValGluLysAsnGluArgLeu 225

Qy 523 CCACGTGATAAAATTTACCTAACCCATTGAGAGCCTGATGACTCGTTGTTGCTCTAAA 582  
Db 226 ThrIleProSerSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245  
Qy 583 GATCCTTCCAGCGCCCTTCAATGGAGGAAATTTGAAAAATAATG 627  
Db 246 AspAlaLysLysArgProSerPheLysGlnIleIleSerIleLeu 260

## RESULT 12

US-09-221-528-5  
; Sequence 5, Application US/09221528  
; Patent No. 6190874  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,528  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-528-5

Alignment Scores:  
Pred. No.: 9.61e-37 Length: 455  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 3 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-221-528-5 (1-455)

Qy 4 GAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCCTGC 63  
Db 53 GluAlaGluIleLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 72  
Qy 64 TTGAATCCA-----GTGTGCTTGTGATGGAATATGCTGAAGGGCTCTTTATATAAT 117  
Db 73 LeuGlnProProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92  
Qy 118 GTGCTGCATGCT-----GCTGAACCATTTGCCATATTATATCTGCTGCCCGCAATGAGT 171  
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109  
Qy 172 TGGTGTTCACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTA 231  
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129  
Qy 232 ATTCACAGGAGCCTGAAACCAACCAACTTACTGCTGTTGTCAGGGGGCAGAGTTCTAAA 291  
Db 130 IleHisArgAspLeuLysSerArgAsnValValIleAlaAspGly---ValLeuLys 148  
Qy 292 ATTTGTGATTTTGTACAGCCTGTGCATTCACACACACATGACCAATAACAAG----- 345  
Db 149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167  
Qy 346 GGGAGTGTCTGGATGGACCTGAAGTTTTTGAAGGTAGTAAATACAGTGAATAATGT 405  
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187  
Qy 406 GACGCTTCACGTGGGTATTATCTTTGGAGAGTAAACGCTCGGAAACCCCTTTGAT 465  
Db 188 AspThrTyrSerTyrGlyValValLeuTrpGluMetLeuThrArgGluValProPheLys 207  
Qy 466 GAGATTGTGGCCAGCTTTCCGAAATCATGTGG---GCTGTTTCATATGATGGTACTCGACCA 522

Db 208 GlyLeuGluGly-----LeuGlnValAlaTriPLeuValValGluLysAsnGluArgLeu 225  
QY 523 CCAGTATAAAATTTACCTAAGCCCATTCAGAGCCTGATGACTGCTTGTGGTCTAAA 582  
Db 226 ThrileProSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245  
QY 583 GATCCTTCCAGGCGCCCTTCAATGGAGGAATTTGTGAAAATAATG 627  
Db 246 AspaLalysLysArgProSerPheLysGlnIleleSerIleLeu 260

## RESULT 13

US-09-593-553-5  
; Sequence 5, Application US/09593553  
; Patent No. 6200770  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/593,553  
; CURRENT FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: 09/163,115  
; PRIOR FILING DATE: 1998-09-28  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-593-553-5

Alignment Scores:  
Pred. No.: 9,61e-37 Length: 455  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 3 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-593-553-5 (1-455)

QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTTATGGAGCTGC 63  
Db 53 GluAlaGluIleLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 72  
QY 64 TTGAATCCA-----GTGTGCTGTGATGGAATATGCTGAAGGGGCTCTTATATAAT 117  
Db 73 LeuGluProProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92  
QY 118 GTGCTGCATGGT-----GCTGAACCATTCGCATATTATATCTGCTGCCCAAGCAATGAGT 171  
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109  
QY 172 TGGTGTTCAGTGTTCCTCAAGGAGTGCTTATCTTCACAGCATGCAACCCAAAGCGCTA 231  
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129  
QY 232 ATTCACAGGAGCCTGAAACCACTTACTGCTGCTGCTGAGGGGAGCAGTCTCTAAA 291  
Db 130 IleHisArgAspLeuLysSerArgAsnValIleAlaAlaAspGly---ValLeuLys 148  
QY 292 ATTTGTGATTTTGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAG----- 345  
Db 149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167  
QY 346 GGGAGTCTGCTTGGATGGCCCTGAAGTTTTCAGAGTAGTAGTATACAGTGAATAATGT 405  
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187  
QY 406 GACGCTCTCAGCTGGGTATTATTCTTTGGAGTGTAAACGGCTCGGAACCCCTTTGAT 465  
Db 188 AspThrTyrSerTyrGlyValValLeuIleProGluMetLeuThrArgGluValProPheLys 207  
QY 466 GAGATTGGTGGCCCGAGCTTTCCGAATCATGTGG---GCTGTTTATAATGGTACTCGACCA 522

Db 208 GlyLeuGluGly-----LeuGlnValAlaTriPLeuValValGluLysAsnGluArgLeu 225  
QY 523 CCAGTATAAAATTTACCTAAGCCCATTCAGAGCCTGATGACTGCTTGTGGTCTAAA 582  
Db 226 ThrileProSerCysProArgSerPheAlaGluLeuLeuHisGlnCysTrpGluAla 245  
QY 583 GATCCTTCCAGGCGCCCTTCAATGGAGGAATTTGTGAAAATAATG 627  
Db 246 AspaLalysLysArgProSerPheLysGlnIleleSerIleLeu 260

## RESULT 14

US-09-221-237-5  
; Sequence 5, Application US/09221237  
; Patent No. 6214597  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,237  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-237-5

Alignment Scores:  
Pred. No.: 9,61e-37 Length: 455  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 3 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x US-09-221-237-5 (1-455)

QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTTATGGAGCTGC 63  
Db 53 GluAlaGluIleLeuSerValLeuSerHisArgAsnIleIleGlnPheTyrGlyValIle 72  
QY 64 TTGAATCCA-----GTGTGCTGTGATGGAATATGCTGAAGGGGCTCTTATATAAT 117  
Db 73 LeuGluProProAsnTyrGlyIleValThrGluTyrAlaSerLeuGlySerLeuTyrAsp 92  
QY 118 GTGCTGCATGGT-----GCTGAACCATTCGCATATTATATCTGCTGCCCAAGCAATGAGT 171  
Db 93 TyrIleAsnSerAsnArgSerGluGluMet-----AspMetAspHisIleMetThr 109  
QY 172 TGGTGTTCAGTGTTCCTCAAGGAGTGCTTATCTTCACAGCATGCAACCCAAAGCGCTA 231  
Db 110 TrpAlaThrAspValAlaLysGlyMetHisTyrLeuHisMetGluAlaProValLysVal 129  
QY 232 ATTCACAGGAGCCTGAAACCACTTACTGCTGCTGCTGAGGGGAGCAGTCTCTAAA 291  
Db 130 IleHisArgAspLeuLysSerArgAsnValIleAlaAlaAspGly---ValLeuLys 148  
QY 292 ATTTGTGATTTTGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAG----- 345  
Db 149 IleCysAspPheGly---AlaSerArgPheHisAsnHisThrThrHisMetSerLeuVal 167  
QY 346 GGGAGTCTGCTTGGATGGCCCTGAAGTTTTCAGAGTAGTAGTATACAGTGAATAATGT 405  
Db 168 GlyThrPheProTrpMetAlaProGluValIleGlnSerLeuProValSerGluThrCys 187  
QY 406 GACGCTCTCAGCTGGGTATTATTCTTTGGAGTGTAAACGGCTCGGAACCCCTTTGAT 465  
Db 188 AspThrTyrSerTyrGlyValValLeuIleProGluMetLeuThrArgGluValProPheLys 207



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 06:12:04 ; Search time 1020.61 Seconds  
(without alignments)  
2227.439 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091  
Perfect score: 684  
Sequence: 1 GTAGAGCTTCGAGTATCCCGTGTGAACCATCTCTATATTTGTAAGCTTTATGGAGCC 684

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues 4403344  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA: \*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	684	100.0	1705	14	US-10-283-023-1
2	684	100.0	1705	14	US-10-283-023-3
3	684	100.0	1788	12	US-10-384-743-14
4	684	100.0	1788	13	US-10-158-895-14
5	684	100.0	2656	12	US-10-384-743-3
6	684	100.0	2656	13	US-10-158-895-3
7	65.4	9.6	1428	10	US-09-938-842A-882
8	56.6	8.3	1398	14	US-09-938-842A-633
9	54.6	8.0	1365	14	US-10-106-698-2068
10	54.6	8.0	1365	9	US-09-757-982-6
11	54.6	8.0	2120	9	US-09-757-982-4
12	54.6	8.0	2251	12	US-10-094-749-838
13	52	7.6	1737	10	US-09-862-027-13
14	52	7.6	3160	14	US-10-177-293-306
15	52	7.6	3268	11	US-09-291-417-1
16	52	7.6	3335	12	US-10-353-690-115

17	46.8	6.8	1239	10	US-09-938-842A-1190	Sequence 1190, Ap
18	46.4	6.8	1926	10	US-09-938-842A-86	Sequence 86, Appl
19	46	6.7	3538	12	US-10-288-798-36	Sequence 36, Appl
20	45.2	6.6	2466	12	US-10-171-404A-47	Sequence 47, Appl
21	45	6.6	3111	13	US-10-014-882-1	Sequence 1, Appli
22	45	6.6	3518	13	US-10-014-882-3	Sequence 3, Appli
23	45	6.6	5549	12	US-10-354-358-23	Sequence 23, Appl
24	43.8	6.4	2157	12	US-10-259-165-367	Sequence 367, Appl
25	43.8	6.4	2160	12	US-10-259-165-39	Sequence 39, Appl
26	43.6	6.4	2146	13	US-10-071-766-83	Sequence 83, Appl
27	43.6	6.4	2148	12	US-10-240-965-251	Sequence 251, App
28	43.6	6.4	2198	9	US-09-012-135A-2	Sequence 2, Appli
29	43.6	6.4	2237	12	US-10-240-965-252	Sequence 252, App
30	43.6	6.4	2283	14	US-10-198-846-10281	Sequence 10281, A
31	43.6	6.4	2284	12	US-09-814-353-21674	Sequence 21674, A
32	43.6	6.4	2342	12	US-10-292-408-27	Sequence 27, Appl
33	43.6	6.4	2347	14	US-10-209-324-1	Sequence 1, Appli
34	43.2	6.3	1887	10	US-09-938-842A-52	Sequence 52, Appl
35	43.2	6.3	2211	10	US-09-938-842A-1577	Sequence 1577, Ap
36	42.6	6.2	1857	10	US-09-764-868-220	Sequence 220, App
37	42.6	6.2	1857	11	US-09-764-891-2341	Sequence 2341, Ap
38	42.6	6.2	1857	14	US-10-205-428-202	Sequence 202, App
39	42.4	6.2	266	10	US-09-878-574-7790	Sequence 7790, Ap
40	41.8	6.1	2564	12	US-10-366-288-11	Sequence 11, Appl
41	41.4	6.1	1143	10	US-09-801-368-401	Sequence 401, App
42	40.4	5.9	2440	14	US-10-153-668-435	Sequence 435, App
43	40.4	5.9	3365	14	US-10-153-668-323	Sequence 323, App
44	40.4	5.9	3428	14	US-10-153-668-437	Sequence 437, App
45	40.4	5.9	3761	12	US-09-814-353-20138	Sequence 20138, A

## ALIGNMENTS

### RESULT 1

US-10-283-023-1  
; Sequence 1, Application US/10283023  
; Publication No. US20030091573A1  
; GENERAL INFORMATION:  
; APPLICANT: Carroll, Joseph M.  
; TITLE OF INVENTION: Methods and compositions for the  
; TITLE OF INVENTION: diagnosis and treatment of hematological disorders using  
; TITLE OF INVENTION: 16319  
; FILE REFERENCE: MP101-239P1RM  
; CURRENT APPLICATION NUMBER: US/10/283,023  
; CURRENT FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: 60/335,044  
; PRIOR FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1705  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-283-023-1

Query Match	100.0%	Score	684	DB	14	Length	1705
Best Local Similarity	100.0%	Pred. No.	3.6e-210				
Matches	684	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	GTAGAGCTTCGAGTATCCCGTGTGAACCATCTCTATATTTGTAAGCTTTATGGAGCC	60				
Db	226	GTAGAGCTTCGAGTATCCCGTGTGAACCATCTCTATATTTGTAAGCTTTATGGAGCC	285				
Qy	61	TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTTTTATATAATGTG	120				
Db	286	TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTTTTATATAATGTG	345				
Qy	121	CTGCATGTGTGAACCATCTGCCATATTTATCTGTCGCCACGCAATGAGTGTGTGTTA	180				
Db	346	CTGCATGTGTGAACCATCTGCCATATTTATCTGTCGCCACGCAATGAGTGTGTGTTA	405				
Qy	191	CAGTGTTCACAGGAGTGGCTTAATCTTACAGCATGCAACCAACGCTAATTCACAGG	240				

Db 406 CAGTGTTCCTCCAGGAGTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAGG 465  
QY 241 GACCTGAACACCAACAACTTACTGCTGGTTCAGGGGGGACAGTTCATAAAATTTGTGAT 300  
Db 466 GACCTGAACACCAACAACTTACTGCTGGTTCAGGGGGGACAGTTCATAAAATTTGTGAT 525  
QY 301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATACAAAGGGGAGTGGCTGTGG 360  
Db 526 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATACAAAGGGGAGTGGCTGTGG 585  
QY 361 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAAGAAATGTGACGCTTTCAGCTGG 420  
Db 586 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAAGAAATGTGACGCTTTCAGCTGG 645  
QY 421 GGTATTATCTTTGGGAAGTGAATACCGGTGCGAAACCCCTTTGATGAGATTTGGTGGCCCA 480  
Db 646 GGTATTATCTTTGGGAAGTGAATACCGGTGCGAAACCCCTTTGATGAGATTTGGTGGCCCA 705  
QY 481 GCTTTCGGAATCATGTGGGCTGTTTCAATAGTGAAGAAATGTGACGCTTTCAGCTGG 540  
Db 706 GCTTTCGGAATCATGTGGGCTGTTTCAATAGTGAAGAAATGTGACGCTTTCAGCTGG 765  
QY 541 CTTAAGCCCATTTGAGCCCTGATGACCTGTTTGGTCTAAGATCCTTCCAGCGCCCT 600  
Db 766 CTTAAGCCCATTTGAGCCCTGATGACCTGTTTGGTCTAAGATCCTTCCAGCGCCCT 825  
QY 601 TCAATGAGGAAATTTGMAAATAATGACTCACTGTGATGGGTACTTTCCAGGAGCAGAT 660  
Db 826 TCAATGAGGAAATTTGMAAATAATGACTCACTGTGATGGGTACTTTCCAGGAGCAGAT 885  
QY 661 GAGCCATTACAGTATCCTTGTCTAG 684  
Db 886 GAGCCATTACAGTATCCTTGTCTAG 909

## RESULT 2

US-10-283-023-3  
; Sequence 3, Application US/10283023  
; Publication No. US20030091573A1  
; GENERAL INFORMATION:  
; APPLICANT: Carroll, Joseph M.  
; TITLE OF INVENTION: Methods and compositions for the  
; diagnosis and treatment of hematological disorders using  
; FILE REFERENCE: MEI01-239PIRM  
; CURRENT APPLICATION NUMBER: US/10/283,023  
; PRIOR FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: 60/335,044  
; PRIOR FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1705  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1557)  
US-10-283-023-3

Query Match 100.0%; Score 684; DB 14; Length 1705;  
Best Local Similarity 100.0%; Pred. No. 3.6e-210;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTAGACTTCGGCAGTATCCCTGTGACCATCTCTAATTTGTAAGCTTTATGAGCC 60  
Db 226 GTAGACTTCGGCAGTATCCCTGTGACCATCTCTAATTTGTAAGCTTTATGAGCC 285  
QY 61 TGTTCGAATCCAGTGTCTTGTGATGGAATGCTGAAGGGGCTCTTTATATATG 120  
Db 286 TGTTCGAATCCAGTGTCTTGTGATGGAATGCTGAAGGGGCTCTTTATATATG 345

QY 121 CTGCATGGTGTGAACCATTCGCATATTATATCTGCTGCCACGCAATGAGTTGGTGTTA 180  
Db 346 CTGCATGGTGTGAACCATTCGCATATTATATCTGCTGCCACGCAATGAGTTGGTGTTA 405  
QY 181 CAGTGTTCCTCCAGGAGTGGCTTATCTTTCAGAGCATGCAACCCAAAGCGCTAATTCACAGG 240  
Db 406 CAGTGTTCCTCCAGGAGTGGCTTATCTTTCAGAGCATGCAACCCAAAGCGCTAATTCACAGG 465  
QY 241 GACCTGAACACCAACAACTTACTGCTGGTTCAGGGGGGACAGTTCATAAAATTTGTGAT 300  
Db 466 GACCTGAACACCAACAACTTACTGCTGGTTCAGGGGGGACAGTTCATAAAATTTGTGAT 525  
QY 301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATACAAAGGGGAGTGGCTGTGG 360  
Db 526 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATACAAAGGGGAGTGGCTGTGG 585  
QY 361 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAAGAAATGTGACGCTTTCAGCTGG 420  
Db 586 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAAGAAATGTGACGCTTTCAGCTGG 645  
QY 421 GGTATTATCTTTGGGAAGTGAATACCGGTGCGAAACCCCTTTGATGAGATTTGGTGGCCCA 480  
Db 646 GGTATTATCTTTGGGAAGTGAATACCGGTGCGAAACCCCTTTGATGAGATTTGGTGGCCCA 705  
QY 481 GCTTTCGGAATCATGTGGGCTGTTTCAATAGTGAAGAAATGTGACGCTTTCAGCTGG 540  
Db 706 GCTTTCGGAATCATGTGGGCTGTTTCAATAGTGAAGAAATGTGACGCTTTCAGCTGG 765  
QY 541 CTTAAGCCCATTTGAGCCCTGATGACCTGTTTGGTCTAAGATCCTTCCAGCGCCCT 600  
Db 766 CTTAAGCCCATTTGAGCCCTGATGACCTGTTTGGTCTAAGATCCTTCCAGCGCCCT 825  
QY 601 TCAATGAGGAAATTTGMAAATAATGACTCACTGTGATGGGTACTTTCCAGGAGCAGAT 660  
Db 826 TCAATGAGGAAATTTGMAAATAATGACTCACTGTGATGGGTACTTTCCAGGAGCAGAT 885  
QY 661 GAGCCATTACAGTATCCTTGTCTAG 684  
Db 886 GAGCCATTACAGTATCCTTGTCTAG 909

## RESULT 3

US-10-384-743-14  
; Sequence 14, Application US/10384743  
; Publication No. US2003016228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; PRIOR FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1788  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)...(1776)  
US-10-384-743-14

Query Match 100.0%; Score 684; DB 12; Length 1788;  
Best Local Similarity 100.0%; Pred. No. 3.7e-210;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAGAGCTTCGGCAGTGTATCCCGTGTGAACCATCTTAATATTGTAAAGCTTTATGGAGCC 60  
Db 232 GTAGAGCTTCGGCAGTGTATCCCGTGTGAACCATCTTAATATTGTAAAGCTTTATGGAGCC 291  
QY 61 TGCCTTGAATCCAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
Db 292 TGCCTTGAATCCAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 351  
QY 121 CTGCATGGTGTGTAACCATTCGCCATATTATATCTGCTGCCACGCAATGAGTTGGTGTGTTA 180  
Db 352 CTGCATGGTGTGTAACCATTCGCCATATTATATCTGCTGCCACGCAATGAGTTGGTGTGTTA 411  
QY 181 CAGTGTTCCTCCAAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 240  
Db 412 CAGTGTTCCTCCAAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 471  
QY 241 GACCTGAAACCAACCAACTTACTGCTGCTGAGGGGGAGCAGTTCTAAAAATTTGTGAT 300  
Db 472 GACCTGAAACCAACCAACTTACTGCTGCTGAGGGGGAGCAGTTCTAAAAATTTGTGAT 531  
QY 301 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTTGG 360  
Db 532 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTTGG 591  
QY 361 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATTTACAGTGAAAAATGTGACGTCTTCAGCTGG 420  
Db 592 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATTTACAGTGAAAAATGTGACGTCTTCAGCTGG 651  
QY 421 GGTATTATCTTTGGGAAGTGTAAACGCTCGGAACCCCTTTGATGAGATTGGTGCCCA 480  
Db 652 GGTATTATCTTTGGGAAGTGTAAACGCTCGGAACCCCTTTGATGAGATTGGTGCCCA 711  
QY 481 GCTTTCCGAATCATGTGGCTGTTTCATAATGCTACTCGACCACTGATGATAAAAAATTTA 540  
Db 712 GCTTTCCGAATCATGTGGCTGTTTCATAATGCTACTCGACCACTGATGATAAAAAATTTA 771  
QY 541 CCTAAGCCCATTTGAGAGCTGATGACTGTTGTTGGTCTAAGATCCTTCCAGGCGCCT 600  
Db 772 CCTAAGCCCATTTGAGAGCTGATGACTGTTGTTGGTCTAAGATCCTTCCAGGCGCCT 831  
QY 601 TCAATGGAGGAAATTTGAAATAATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 660  
Db 832 TCAATGGAGGAAATTTGAAATAATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 891  
QY 661 GAGCCATTACAGTATCCTTTGTGAG 684  
Db 892 GAGCCATTACAGTATCCTTTGTGAG 915

## RESULT 4

US-10-158-895-14  
; Sequence 14, Application US/10158895  
; Publication No. US2002015624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1788  
; TYPE: DNA

; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1776)  
US-10-158-895-14

Query Match 100.0%; Score 684; DB 13; Length 1788;  
Best Local Similarity 100.0%; Pred. No. 3.7e-210; Indels 0; Gaps 0;  
Matches 684; Conservative 0; Mismatches 0;

QY 1 GTAGAGCTTCGGCAGTGTATCCCGTGTGAACCATCTTAATATTGTAAAGCTTTATGGAGCC 60  
Db 232 GTAGAGCTTCGGCAGTGTATCCCGTGTGAACCATCTTAATATTGTAAAGCTTTATGGAGCC 291  
QY 61 TGCCTTGAATCCAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
Db 292 TGCCTTGAATCCAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 351  
QY 121 CTGCATGGTGTGTAACCATTCGCCATATTATATCTGCTGCCACGCAATGAGTTGGTGTGTTA 180  
Db 352 CTGCATGGTGTGTAACCATTCGCCATATTATATCTGCTGCCACGCAATGAGTTGGTGTGTTA 411  
QY 181 CAGTGTTCCTCCAAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 240  
Db 412 CAGTGTTCCTCCAAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 471  
QY 241 GACCTGAAACCAACCAACTTACTGCTGCTGAGGGGGAGCAGTTCTAAAAATTTGTGAT 300  
Db 472 GACCTGAAACCAACCAACTTACTGCTGCTGAGGGGGAGCAGTTCTAAAAATTTGTGAT 531  
QY 301 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTTGG 360  
Db 532 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTTGG 591  
QY 361 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATTTACAGTGAAAAATGTGACGTCTTCAGCTGG 420  
Db 592 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATTTACAGTGAAAAATGTGACGTCTTCAGCTGG 651  
QY 421 GGTATTATCTTTGGGAAGTGTAAACGCTCGGAACCCCTTTGATGAGATTGGTGCCCA 480  
Db 652 GGTATTATCTTTGGGAAGTGTAAACGCTCGGAACCCCTTTGATGAGATTGGTGCCCA 711  
QY 481 GCTTTCCGAATCATGTGGCTGTTTCATAATGCTACTCGACCACTGATGATAAAAAATTTA 540  
Db 712 GCTTTCCGAATCATGTGGCTGTTTCATAATGCTACTCGACCACTGATGATAAAAAATTTA 771  
QY 541 CCTAAGCCCATTTGAGAGCTGATGACTGTTGTTGGTCTAAGATCCTTCCAGGCGCCT 600  
Db 772 CCTAAGCCCATTTGAGAGCTGATGACTGTTGTTGGTCTAAGATCCTTCCAGGCGCCT 831  
QY 601 TCAATGGAGGAAATTTGAAATAATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 660  
Db 832 TCAATGGAGGAAATTTGAAATAATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 891  
QY 661 GAGCCATTACAGTATCCTTTGTGAG 684  
Db 892 GAGCCATTACAGTATCCTTTGTGAG 915

## RESULT 5

US-10-384-743-3  
; Sequence 3, Application US/10384743  
; Publication No. US2003016228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; FILE REFERENCE: 053466/0278  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11

; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2656  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (183)..(1919)  
US-10-384-743-3

Query Match 100.0%; Score 684; DB 12; Length 2656;  
Best Local Similarity 100.0%; Pred. No. 4.6e-210;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 60  
Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 467  
QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATCTGAAGGGGCTCTTTATATAAATGTG 120  
Db 468 TGCTTGAATCCAGTGTCTTGTGATGGAATATCTGAAGGGGCTCTTTATATAAATGTG 527  
QY 121 CTGCATGGTGTGAACCATCTCCCATATTAATCTGCTGCCACGCAATGAGTGTGTGTTA 180  
Db 528 CTGCATGGTGTGAACCATCTCCCATATTAATCTGCTGCCACGCAATGAGTGTGTGTTA 587  
QY 181 CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGG 240  
Db 588 CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGG 647  
QY 241 GACCTGAAACCCACCAACTTACTGCTGGTTCAGGGGGGACAGTTCTTAAATAATTTGTGAT 300  
Db 648 GACCTGAAACCCACCAACTTACTGCTGGTTCAGGGGGGACAGTTCTTAAATAATTTGTGAT 707  
QY 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGCTTGG 360  
Db 708 TTTGGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGCTTGG 767  
QY 361 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAATAATCTGACGCTCTTCAGCTGG 420  
Db 768 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAATAATCTGACGCTCTTCAGCTGG 827  
QY 421 GGTATTAATCTTTGGGAAGTGAATTAACCGTGGGAAACCCCTTTGATGAGATTGGTGGCCCA 480  
Db 828 GGTATTAATCTTTGGGAAGTGAATTAACCGTGGGAAACCCCTTTGATGAGATTGGTGGCCCA 887  
QY 481 GCTTTCGGAATCATGTGGGCTGTTTCAATGTTACTGACACCACTGATGATAAAAAATTTA 540  
Db 888 GCTTTCGGAATCATGTGGGCTGTTTCAATGTTACTGACACCACTGATGATAAAAAATTTA 947  
QY 541 CCTAAGCCCATTTGAGAGCCTGATGACTCGTGTGTGGTCTAAAGATCTCTCCAGCGCCT 600  
Db 948 CCTAAGCCCATTTGAGAGCCTGATGACTCGTGTGTGGTCTAAAGATCTCTCCAGCGCCT 1007  
QY 601 TCAATGAGGAAATTTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 660  
Db 1008 TCAATGAGGAAATTTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 1067  
QY 661 GAGCCATTACAGTATCCTTTGTCAG 684  
Db 1068 GAGCCATTACAGTATCCTTTGTCAG 1091

RESULT 6  
US-10-158-895-3  
; Sequence 3, Application US/1015895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:

; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2656  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (183)..(1919)  
US-10-158-895-3

Query Match 100.0%; Score 684; DB 13; Length 2656;  
Best Local Similarity 100.0%; Pred. No. 4.6e-210;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 60  
Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 467  
QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATCTGAAGGGGCTCTTTATATAAATGTG 120  
Db 468 TGCTTGAATCCAGTGTCTTGTGATGGAATATCTGAAGGGGCTCTTTATATAAATGTG 527  
QY 121 CTGCATGGTGTGAACCATCTCCCATATTAATCTGCTGCCACGCAATGAGTGTGTGTTA 180  
Db 528 CTGCATGGTGTGAACCATCTCCCATATTAATCTGCTGCCACGCAATGAGTGTGTGTTA 587  
QY 181 CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGG 240  
Db 588 CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGG 647  
QY 241 GACCTGAAACCCACCAACTTACTGCTGGTTCAGGGGGGACAGTTCTTAAATAATTTGTGAT 300  
Db 648 GACCTGAAACCCACCAACTTACTGCTGGTTCAGGGGGGACAGTTCTTAAATAATTTGTGAT 707  
QY 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGCTTGG 360  
Db 708 TTTGGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGCTTGG 767  
QY 361 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAATAATCTGACGCTCTTCAGCTGG 420  
Db 768 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAATAATCTGACGCTCTTCAGCTGG 827  
QY 421 GGTATTAATCTTTGGGAAGTGAATTAACCGTGGGAAACCCCTTTGATGAGATTGGTGGCCCA 480  
Db 828 GGTATTAATCTTTGGGAAGTGAATTAACCGTGGGAAACCCCTTTGATGAGATTGGTGGCCCA 887  
QY 481 GCTTTCGGAATCATGTGGGCTGTTTCAATGTTACTGACACCACTGATGATAAAAAATTTA 540  
Db 888 GCTTTCGGAATCATGTGGGCTGTTTCAATGTTACTGACACCACTGATGATAAAAAATTTA 947  
QY 541 CCTAAGCCCATTTGAGAGCCTGATGACTCGTGTGTGGTCTAAAGATCTCTCCAGCGCCT 600  
Db 948 CCTAAGCCCATTTGAGAGCCTGATGACTCGTGTGTGGTCTAAAGATCTCTCCAGCGCCT 1007  
QY 601 TCAATGAGGAAATTTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 660  
Db 1008 TCAATGAGGAAATTTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 1067  
QY 661 GAGCCATTACAGTATCCTTTGTCAG 684



```
Db 1068 GAGCATTACAGTATCCTTGTCAG 1091
|||||
RESULT 7
US-09-938-842A-882
; Sequence 882, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 882
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-882

Query Match 9.6%; Score 65.4; DB 10; Length 1428;
Best Local Similarity 49.3%; Pred. No. 3.9e-10;
Matches 171; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 338 ATACCAAGGGAGTCTGCTTGGATGGCACCCTGAAAGTTTTTGAAGGTAGTAATTACAGTG 397
|||||
Db 956 ATACATAGGAACCTTATAGGTGGATGGCACCCTGAAAGTTTTTAAACGGATACCAATGAC 1015
|||||
QY 398 AAAAATGTGACGTCTTCAGCTGGGGTATATTCTTTGGGAAGTGATAACGGCTCGGAAC 457
|||||
Db 1016 GGAAGTGGCATGTTTATAGTTTGGACTTCTTTATGGMAATGTTAGTGGAGCACTTC 1075
|||||
QY 458 CTTTGTATGAGATTGGTGGCCAGCTTTCCGNAATCATGTGGGCTGTTTATATGATCTC 517
|||||
Db 1076 CATATGAGGAGATGAAATTTGCTGAACAAATTCCTACGAGTTATATACAGAAATTA 1135
|||||
QY 518 GACCACCACTGATAAAAAATTTACCTAAGCCATTGAGACCTGATGACTGCTTGTGTT 577
|||||
Db 1136 GGCCAGTTATACCGACGAGTTGTCAGCGGCCATGAAGAGCTGATCGAGCATTTGGT 1195
|||||
QY 578 CTAAGATCTCTCCAGCGCCCTTCAATGGAGGAAATTTGGAATAATGATCACTTGA 637
|||||
Db 1196 CATCGAAACACAGACAGAGACCGGAATTTCTGGCAGATTGTCAAAGTGTGGAACATTCA 1255
|||||
QY 638 TCGGTACTTTCCAGGACAGATGAGCCATTACAGTATCCTTGTCTAG 684
|||||
Db 1256 AGAAGTCTCTAACAGCGAAGAAAACCTTAATCTTTTACCTAGCCAG 1302
|||||

RESULT 8
US-09-938-842A-633
; Sequence 633, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
```

```
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 633
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-633

Query Match 8.3%; Score 56.6; DB 10; Length 1398;
Best Local Similarity 52.3%; Pred. No. 2.7e-07;
Matches 150; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 186 TTCCCAAGGAGTGGCTTATCTTCACAGCATCAACCCAAAGCGCTAATTTCACAGGACCT 245
|||||
Db 714 TCGAATAATCGTCTCTGCGATTTCCCATCTCCATGAGAAAGGCATATATGCACAGATCT 773
|||||
QY 246 GAAACCAACAAACTTACTGCTGGTTGAGGGGGGACAG---TTCUAAAAATTTGTGATTT 302
|||||
Db 774 GAACCCGAAACATACTCATGATACAGATGGCCACGTGATGCTAACAGATTTTGGTT 833
|||||
QY 303 TGGTACAGCCTGTGACATTCAGACACATGACCAATTAACAGGGAGTGTGCTTGGAT 362
|||||
Db 834 AGCAAGGAATTTGAAGAAACACACAAGATCAAACTCCATGTGCGGAACCTACCGAGTATAT 893
|||||
QY 363 GGCACCTGAAGTTTTTGAAGGTAGTAATTACAGTGAAATAATGTACGCTTTCAGCTGGG 422
|||||
Db 894 GGCACCTGAATGTTTCGAGGAAAGGACATGATAAGCAGCTGACTGCTGGTGAGCGTAGG 953
|||||
QY 423 TATTATCTTTGGGAAGTGATAACGCGTCGGAAACCCCTTTTGATGAGA 469
|||||
Db 954 GATCTCTGTATGAGATGCTCACAGGAAGCCACCGTTTCTCGGA 1000
|||||

RESULT 9
US-10-106-698-2068/c
; Sequence 2068, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent in Ver. 3.0
; SEQ ID NO 2068
; LENGTH: 1063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-2068

Query Match 8.0%; Score 54.6; DB 14; Length 1063;
Best Local Similarity 48.3%; Pred. No. 1e-06;
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 160 CACGCAATGAGTTGGTGTTTACAGTTTCCCAAGAGTGGCTTATCTTCACAGCATGCAA 219
|||||
Db 578 CACATTATGACCTGGCCCATGATGTAGCCAAAGAAATGCAATTATTACATATGGAGCT 519
|||||
QY 220 CCCAAAGCGCTAATTTCACAGGACCTGAAACCCAACTTACTCTGCTGTTGCAGGGGG 279
|||||
```

Db 518 CCTGTCAAGGTGATTACACAGAGACCTCAAGTCAAGAAACGGTTGTTATAGCTGTGATGGA 459  
QY 280 ACAGTTCCTAAATTTGTGATTTTGGTACAGCCTGTGACATTCACACACATCAACCAAT 339  
Db 458 GTATTGAAGATCTGTGATTTGGTGGCTCTCGGTTCCATACCAATCAACACACATGTCC 399  
QY 340 AACAAAGGGAGTGTCTGTGGATGGCACCCTGAAGTTTGAAGTGTAGTAATTACAGTGAA 399  
Db 398 TTGGTTGAACCTTCCCATGATGGCTCCAGAAAGTTATCCAGATCTCCCTGTGTGAGAA 339  
QY 400 AAATGTCAAGCTTTCAAGCTGGGTATTATCTTTGGAAAGTGAATAACGGCTCGGAAACCC 459  
Db 338 ACTGTGACACATATTCATAGTGTGGTCTCTGGGAGATGCTAAACAAGGAGGTCCCC 279  
QY 460 TTTGATGAGATTGGTG 476  
Db 278 TTTAAAGGTTTGAAGG 262

## RESULT 10

US-09-757-982-6

; Sequence 6, Application US/09757982  
; Patent No. US20020094559A1

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/757,982

; CURRENT FILING DATE: 2001-01-10

; PRIOR APPLICATION NUMBER: 09/163,115

; PRIOR FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1365

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)...(1365)

; US-09-757-982-6

Query Match 8.0%; Score 54.6; DB 9; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 1.2e-06;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 160 CAGCAATGAGTTGGTGTTCACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAA 219  
Db 316 CACATTATGACCTGGGCCACTGTATGAGCCAAAGGAATGCATTATTACATATGGAGCT 375  
QY 220 CCCAAAGCGCTAATTCACAGGACCTGAACCCAACTTACTGCTGGTTGACAGGGGG 279  
Db 376 CCTGTCAAGGTGATTCACAGAGACCTCAAGTCAAGAAACGGTTGTTATAGCTGTGATGGA 435  
QY 280 ACAGTTCCTAAATTTGTGATTTTGGTACAGCCTGTGACATTCACACACATCAACCAAT 339  
Db 436 GTACTGAAGATCTGTGATTTGGTGGCTCTCGGTTCCATACCAATCAACACATGTCC 495  
QY 340 AACAAAGGGAGTGTCTGTGGATGGCACCCTGAAGTTTGAAGTGTAGTAATTACAGTGAA 399  
Db 496 TTGGTTGAACCTTCCCATGATGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTGAGAA 555  
QY 400 AAATGTCAAGCTTTCAAGCTGGGTATTATCTTTGGAAAGTGAATAACGGCTCGGAAACCC 459  
Db 556 ACTGTGACACATATTCATAGTGTGGTCTCTGGGAGATGCTAAACAAGGAGGTCCCC 615  
QY 460 TTTGATGAGATTGGTG 476  
Db 616 TTTAAAGGTTTGAAGG 632

## RESULT 11

US-09-757-982-4

; Sequence 4, Application US/09757982  
; Patent No. US20020094559A1

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/757,982

; CURRENT FILING DATE: 2001-01-10

; PRIOR APPLICATION NUMBER: 09/163,115

; PRIOR FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 2120

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (47)...(1411)

; US-09-757-982-4

Query Match 8.0%; Score 54.6; DB 9; Length 2120;  
Best Local Similarity 48.3%; Pred. No. 1.5e-06;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 160 CAGCAATGAGTTGGTGTTCACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAA 219  
Db 362 CACATTATGACCTGGGCCACTGTATGAGCCAAAGGAATGCATTATTACATATGGAGCT 421  
QY 220 CCCAAAGCGCTAATTCACAGGACCTGAACCCAACTTACTGCTGGTTGACAGGGGG 279  
Db 422 CCTGTCAAGGTGATTCACAGAGACCTCAAGTCAAGAAACGGTTGTTATAGCTGTGATGGA 481  
QY 280 ACAGTTCCTAAATTTGTGATTTTGGTACAGCCTGTGACATTCACACACATCAACCAAT 339  
Db 482 GTACTGAAGATCTGTGACTTTGGTGGCTCTCGGTTCCATACCAATCAACACATGTCC 541  
QY 340 AACAAAGGGAGTGTCTGTGGATGGCACCCTGAAGTTTGAAGTGTAGTAATTACAGTGAA 399  
Db 542 TTGGTTGAACCTTCCCATGATGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTGAGAA 601  
QY 400 AAATGTCAAGCTTTCAAGCTGGGTATTATCTTTGGAAAGTGAATAACGGCTCGGAAACCC 459  
Db 602 ACTGTGACACATATTCATAGTGTGGTCTCTGGGAGATGCTAAACAAGGAGGTCCCC 661  
QY 460 TTTGATGAGATTGGTG 476  
Db 662 TTTAAAGGTTTGAAGG 678

## RESULT 12

US-10-094-749-838

; Sequence 838, Application US/10094749

; Publication No. US20030219741A1

; GENERAL INFORMATION:

; APPLICANT: ISOGAI, TAKAO

; APPLICANT: SUGIYAMA, TOMOYASU

; APPLICANT: OTSUKI, TETSUJI

; APPLICANT: WAKAMATSU, AI

; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO

; APPLICANT: YAMAMOTO, JUN-ICHI

; APPLICANT: ISONO, YUKIO

; APPLICANT: HIO, YURI

; APPLICANT: OTSUKA, KAORU

; APPLICANT: NAGAI, KEIICHI

; APPLICANT: IRIE, RYOTARO

; APPLICANT: TAMECHIKA, ICHIRO

; APPLICANT: SEKI, NAOHICO

; APPLICANT: YOSHIKAWA, TSUTOMU

; APPLICANT: OTSUKA, MOTOKYUKI

; APPLICANT: NAGAHARI, KENJI

; APPLICANT: MASUHO, YASUHIKO

```
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; BEST FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 838
; LENGTH: 2251
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-094-749-838

Query Match      8.0%; Score 54.6; DB 12; Length 2251;
Best Local Similarity 48.3%; Pred. No. 1.6e-06;
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 160 CACGAATGAGTGGTGTACAGTGTCCCAAGAGTGGCTATCTTCACAGCATGCAA 219
DB 517 CACATTATGACCTGGGCCACTGATAGCAAGGAATGCATTATTACATATGGAGGCT 576
QY 220 CCCAAGCGCTAAATTCACAGGACCTGAAACCAACAACTTACTGCTGTTGCAGGGGG 279
DB 577 CCTGTCAAGGTGATTCACAGAGACCTCAAGTCAAGAAACGTTGTATAGCTGCTGATGA 636
QY 280 ACAGTCTTAAAAATTTGATTTTGGTACAGCTGTGACATTCACAGACACATGACCAAT 339
DB 637 GTATTGAAGATCTGAGCTTTGGTGGCTCTCGTTCCATACCAACACACATGTC 696
QY 340 AACAGGGGAGTGTGCTGATGAGTGCACCTGAAGTTTGAAGTAGTAAATACAGTGA 399
DB 697 TTGGTTGGAACCTTCCCATGATGCTCCAGAGTTATCCAGAGTCTCCCTGTGTCAAG 756
QY 400 AAATGTGACGCTTCAGCTGGGTATTTCTTTCGGGAAGTGAATACCGCTCGGAAACC 459
DB 757 ACTTGTGACACATATTCCTATGTTGGTGTCTCTGGAGATGCTAAACAGGAGTCC 816
QY 460 TTTGATGAGATTGGTGG 476
DB 817 TTTAAGGTTTGAAGG 833

RESULT 13
US-09-862-027-13
; Sequence 13, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: NO. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (275) ... (1522)
US-09-862-027-13

Query Match      7.6%; Score 52; DB 10; Length 1737;
Best Local Similarity 48.8%; Pred. No. 9.5e-06;
Matches 207; Conservative 0; Mismatches 205; Indels 12; Gaps 2;

QY 204 TCTTCACAGCATGCAACCAAGCGCTAATTCACAGGAGCTGAAACCAACCAACTTACT 263
```

```
DB 667 TCTGCACTATCTGCATTCAAGAAAGAAAATTCACGACATAAAAGCTGCCATGTCTT 726
QY 264 GCT-----GGTGCAGGGGGGACAGTTCTAAAAATTTGTGATTTTGTGTACAGCCTGTGA 317
DB 727 GCTCTCAGAACACAGGAGATGTTAAACTTGTCTGATTTTGGAGTTGCTGCTCAGCTCACAGA 786
QY 318 CATTCAGACACATGACCAATAACAAGGGGAGTCTGCTTGGATGGCACCTGAAGTTT 377
DB 787 TACACAGATTAAAAAATAACCTTTGTGGGAACCTCACTTTTGGATGGCTCTCTGAAGTTAT 846
QY 378 TGAAGGTAGTAATTACAGTGAATAATGTACGCTCTTACGCTGGGTATTTCTTTGGGA 437
DB 847 TCAACAGTCAGCTTATGACTCAAAAGCTGACATTTGGTCATTTGGGAATTTACTGCTATTGA 906
QY 438 AGTGATAACGCTCGGAACCCCTTTGATGATGATGGTGGCCCGCAGCTTTCCGAATCATGTG 497
DB 907 ACTAGCCAAAGGAGAGCCACCTAACTCCGATATGCATC-----CAATGAGAGTTCTGTT 960
QY 498 GCCTGTTCAATATGGTACTCGACCACTGATAAAAAATTTACCTAGCCCATTTGAGAG 557
DB 961 TCTTATTTCCCAAAACAATCTCCAACTCTTTGTTGGAGACTTTACTAAGTCTTTTAAAGA 1020
QY 558 CCGTATGACTCGTTGGTCTTAAAGATCCCTCCAGCGCCCTTCAATGGAGGAAATTTGT 617
DB 1021 GTTATTGATGCTTGGCTGAACAAAGATCCATCATCTTCTCTACAGCAAAAGAACTTCT 1080
QY 618 GAAA 621
DB 1081 GAAA 1084

RESULT 14
US-10-177-293-306
; Sequence 306, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
```

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 306

; LENGTH: 3160

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-177-293-306

Query Match

7.6%; Score 52; DB 14; Length 3160;

Best Local Similarity 48.8%; Pred. No. 1.3e-05;

Matches 207; Conservative 0; Mismatches 205; Indels 12; Gaps 2;

Qy 204 TCTTCACAGCATGCAACCCAAAGCGCTAAATTCACAGGAGCTGAAACCAACCAACTTACT 263

Db 510 TCTGGACTATCTGCATTTCAGAAAAAGAAATTCACCGAGACATAAAAGCTGCCAATGCTT 569

Qy 264 GCT-----GGTTCCAGGGGGACAGTCTTAAATAATTGTGATTTTGGTACAGCCTGTGA 317

Db 570 GCTCTCAGAACAGGAGATGTTAAACTTGTGATTTTGGAGTTGCTGGTACAGCTGACAGA 629

Qy 318 CATTTCAGACACACATGACCAATAACAAGGGGAGTGTCTTGGATGGCACCCTGAAGTTT 377

Db 630 TACACAGATTAAAGAAATACCTTTGTGGAACTCCATTTTGGATGGCTCTCTGAAGTTAT 689

Qy 378 TGAAGGTAGTAATTACAGTGAATAATGTGAGCTTTCAGCTGGGGTATATCTTTGGGA 437

Db 690 TCAACAGTCAGCTTATGACTCAAAAGAGCTGACATTTGGTTCATTTGGGAATTACTGCTATTGA 749

Qy 438 AGTGATAACGGTCCGGAACCTTTGATGAGATTGGTGGCCAGCTTCCGAATCATGTG 497

Db 750 ACTAGCCAAGGAGAGCCACTAACTCCGATATGCATC-----CAATGAGAGTCTCTTT 803

Qy 498 GGCTGTTCAVTAATGGTACTGCACCACCACTGATAAAAAATTTACCTAAGCCCATTTGAGAG 557

Db 804 TCTTATCCCAAAACAATCTCCAACTCTTGTGGAGACTTTACTAAGTCTTTTAAGGA 863

Qy 558 CCTGATGACTCGTTTGTGTTAAAGATCCCTTCCAGCGCCCTTCAATGGAGGAAATGT 617

Db 864 GTTTATTGATGCTTGGCTGGAACAAAGATCCATCATTTTCGTCTACAGCAAAAGAACTTCT 923

Qy 618 GAAA 621

Db 924 GAAA 927

RESULT 15

US-09-291-417-1

; Sequence 1, Application US/09291417A

; Publication No. US20030050230A1

; GENERAL INFORMATION:

; APPLICANT: PLOWMAN, GREGORY

; APPLICANT: MARTINEZ, RICARDO

; APPLICANT: WHITE, DAVID

; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES

; FILE REFERENCE: 240/300

; CURRENT APPLICATION NUMBER: US/09/291.417A

; CURRENT FILING DATE: 1999-04-13

; EARLIER APPLICATION NUMBER: US 60/081,784

; EARLIER FILING DATE: 1998-04-14

; NUMBER OF SEQ ID NOS: 147

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 3268

; TYPE: DNA

; ORGANISM: Mammalian (Human) STUK2

Query Match

7.6%; Score 52; DB 11; Length 3268;

Best Local Similarity 48.8%; Pred. No. 1.3e-05;

Matches 207; Conservative 0; Mismatches 205; Indels 12; Gaps 2;

Qy 204 TCTTCACAGCATGCAACCCAAAGCGCTAAATTCACAGGAGCTGAAACCAACCAACTTACT 263

Db 574 TCTGGACTATCTGCATTTCAGAAAAAGAAATTCACCGAGACATAAAAGCTGCCAATGCTT 633

Qy 264 GCT-----GGTTCCAGGGGGACAGTCTTAAATAATTGTGATTTTGGTACAGCCTGTGA 317

Db 634 GCTCTCAGAACAGGAGATGTTAAACTTGTGATTTTGGAGTTGCTGGTACAGCTGACAGA 693

Qy 318 CATTTCAGACACACATGACCAATAACAAGGGGAGTGTCTTGGATGGCACCCTGAAGTTT 377

Db 694 TACACAGATTAAAGAAATACCTTTGTGGAACTCCATTTTGGATGGCTCTCTGAAGTTAT 753

Qy 378 TGAAGGTAGTAATTACAGTGAATAATGTGAGCTTTCAGCTGGGGTATATCTTTGGGA 437

Db 754 TCAACAGTCAGCTTATGACTCAAAAGAGCTGACATTTGGTTCATTTGGGAATTACTGCTATTGA 813

Qy 438 AGTGATAACGGTCCGGAACCTTTGATGAGATTGGTGGCCAGCTTCCGAATCATGTG 497

Db 814 ACTAGCCAAGGAGAGCCACTAACTCCGATATGCATC-----CAATGAGAGTCTCTTT 867

Qy 498 GGCTGTTCAVTAATGGTACTGCACCACCACTGATAAAAAATTTACCTAAGCCCATTTGAGAG 557

Db 868 TCTTATCCCAAAACAATCTCCAACTCTTGTGGAGACTTTACTAAGTCTTTTAAGGA 927

Qy 558 CCTGATGACTCGTTTGTGTTAAAGATCCCTTCCAGCGCCCTTCAATGGAGGAAATGT 617

Db 928 GTTTATTGATGCTTGGCTGGAACAAAGATCCATCATTTTCGTCTACAGCAAAAGAACTTCT 987

Qy 618 GAAA 621

Db 988 GAAA 991

Search completed: December 4, 2003, 09:00:00.

Job time : 1022.11 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:33:11 ; Search time 70.8649 Seconds  
(without alignments)  
598.382 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303

Perfect score: 1252

Sequence: 1 VELQLSRVNHPIVYKLYGA.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1252	100.0	518	15	US-10-283-023-2
2	1252	100.0	579	14	US-10-384-743-4
3	1252	100.0	579	14	US-10-158-895-4
4	1252	100.0	590	12	US-10-384-743-15
5	1252	100.0	590	14	US-10-158-895-15
6	371	29.6	349	15	US-10-106-698-6345
7	371	29.6	455	9	US-09-757-982-5
8	371	29.6	455	12	US-10-094-749-2477
9	363.5	29.0	394	10	US-09-862-027-19
10	355	28.4	1036	12	US-10-354-358-24
11	355	28.4	1036	14	US-10-014-882-2
12	353.5	28.2	746	15	US-10-153-668-436
13	353.5	28.2	859	15	US-10-153-668-324
14	353.5	28.2	892	15	US-10-153-668-438
15	349.5	27.9	1097	12	US-10-288-798-12

16	346	27.6	328	10	US-09-862-027-18	Sequence 18, Appl
17	345.5	27.6	847	14	US-10-143-133-2	Sequence 2, Appl
18	341	27.2	252	12	US-09-976-782-41	Sequence 41, Appl
19	341	27.2	254	12	US-09-976-782-30	Sequence 30, Appl
20	341	27.2	256	12	US-09-863-776-41	Sequence 41, Appl
21	338.5	27.0	835	10	US-09-947-199-2	Sequence 2, Appl
22	335.5	26.8	835	10	US-09-947-199-8	Sequence 8, Appl
23	331.5	26.5	966	10	US-09-771-161A-197	Sequence 197, Appl
24	326.5	26.1	256	12	US-09-976-782-40	Sequence 40, Appl
25	326.5	26.1	257	12	US-09-976-782-29	Sequence 29, Appl
26	323	25.8	257	11	US-09-823-187-46	Sequence 46, Appl
27	323	25.8	257	12	US-09-863-776-42	Sequence 42, Appl
28	323	25.8	821	12	US-10-171-404A-48	Sequence 48, Appl
29	320	25.6	263	10	US-09-840-704-5	Sequence 5, Appl
30	313.5	25.0	850	10	US-09-904-389-2	Sequence 2, Appl
31	312	24.9	265	9	US-09-797-039-10	Sequence 10, Appl
32	312	24.9	265	12	US-10-170-789-10	Sequence 10, Appl
33	312	24.9	277	9	US-09-815-915-13	Sequence 13, Appl
34	312	24.9	277	10	US-09-882-166-4	Sequence 4, Appl
35	312	24.9	277	12	US-10-393-316-13	Sequence 13, Appl
36	312	24.9	277	12	US-10-170-789-17	Sequence 17, Appl
37	312	24.9	277	15	US-10-172-088-6	Sequence 6, Appl
38	312	24.9	278	9	US-09-797-039-13	Sequence 13, Appl
39	312	24.9	278	9	US-09-922-138-18	Sequence 18, Appl
40	312	24.9	278	9	US-09-922-138-27	Sequence 27, Appl
41	312	24.9	278	9	US-09-910-150-17	Sequence 17, Appl
42	312	24.9	278	9	US-09-910-150-31	Sequence 31, Appl
43	312	24.9	278	10	US-09-842-582-4	Sequence 4, Appl
44	312	24.9	278	10	US-09-934-406-4	Sequence 4, Appl
45	312	24.9	278	12	US-10-170-789-13	Sequence 13, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-283-023-2  
; Sequence 2, Application US/10283023  
; Publication No. US20030091573A1  
; GENERAL INFORMATION:  
; APPLICANT: Caroll, Joseph M.  
; TITLE OF INVENTION: Methods and compositions for the  
; TITLE OF INVENTION: diagnosis and treatment of hematological disorders using  
; TITLE OF INVENTION: 16313  
; FILE REFERENCE: MP101-239P1RM  
; CURRENT APPLICATION NUMBER: US/10/283,023  
; CURRENT FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: 60/335,044  
; PRIOR FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-283-023-2

Query Match 100.0%; Score 1252; DB 15; Length 518;  
Best Local Similarity 100.0%; Pred. No. 9.2e-113;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VELQLSRVNHPIVYKLYGACLNIPVCLVMEYAEGLSNVLHGAEPLPYTTAAHMSWCL	60
Db	76	VELQLSRVNHPIVYKLYGACLNIPVCLVMEYAEGLSNVLHGAEPLPYTTAAHMSWCL	135
Qy	61	QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLKICDFGTACDIOITHNNKGSAAW	120
Db	136	QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLKICDFGTACDIOITHNNKGSAAW	195
Qy	121	MAPEFEGSNYSKCDVFSWGIIWEVITRRKPFDEIGGPAPRIMWAVHNGTRPPLIKNL	180
Db	196	MAPEFEGSNYSKCDVFSWGIIWEVITRRKPFDEIGGPAPRIMWAVHNGTRPPLIKNL	255

QY 181 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 228  
Db 256 FKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 303

## RESULT 2

US-10-384-743-4

; Sequence 4, Application US/10384743  
; Publication No. US20030162228A1

; GENERAL INFORMATION:

; APPLICANT: OHTOMO, KOICHIRO

; APPLICANT: TSUCHIYA, MASAYUKI

; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278

; CURRENT APPLICATION NUMBER: US/10/384,743

; CURRENT FILING DATE: 2003-03-11

; PRIOR APPLICATION NUMBER: US/09/529,279

; PRIOR FILING DATE: 2000-04-11

; PRIOR APPLICATION NUMBER: PCT/JP98/04796

; PRIOR FILING DATE: 1998-10-22

; PRIOR APPLICATION NUMBER: JP 9/290188

; PRIOR FILING DATE: 1997-10-22

; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 4

; LENGTH: 579

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-384-743-4

Query Match 100.0%; Score 1252; DB 12; Length 579;

Best Local Similarity 100.0%; Pred. No. 1.1e-112;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 60

Db 76 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 135

QY 61 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 120

Db 136 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 195

QY 121 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180

Db 196 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 228

Db 256 FKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 303

US-10-384-743-15

Query Match 100.0%; Score 1252; DB 12; Length 590;

Best Local Similarity 100.0%; Pred. No. 1.1e-112;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 60

Db 76 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 135

QY 61 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 120

Db 136 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 195

QY 121 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180

Db 196 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 228

Db 256 FKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 303

US-10-384-743-15

Query Match 100.0%; Score 1252; DB 12; Length 590;

Best Local Similarity 100.0%; Pred. No. 1.1e-112;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 60

Db 76 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 135

QY 61 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 120

Db 136 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 195

QY 121 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180

Db 196 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 228

Db 256 FKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 303

US-10-384-743-15

Query Match 100.0%; Score 1252; DB 12; Length 590;

Best Local Similarity 100.0%; Pred. No. 1.1e-112;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 60

Db 76 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 135

QY 61 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 120

Db 136 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 195

QY 121 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180

Db 196 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 228

Db 256 FKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 303

US-10-384-743-15

Query Match 100.0%; Score 1252; DB 12; Length 590;

Best Local Similarity 100.0%; Pred. No. 1.1e-112;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 60

Db 76 VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAEPPLPYTAAHAMSACL 135

QY 61 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 120

Db 136 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDPGTACDIQTHMTNNKGSAAW 195

QY 121 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180

Db 196 MAPEVFEAGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 228

Db 256 FKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADPELOYPCQ 303

US-10-384-743-15

```

RESULT 5
US-10-158-895-15
; Sequence 15, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-15

Query Match      100.0%; Score 1252; DB 14; Length 590;
Best Local Similarity 100.0%; Pred. No. 1.1e-112;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHNPVIVKLYGACLNPPVCLVMEYAEGLSYNNVLHGAEPLPYTAAHAMSACL 60
DB 76 VELRLSRVNHNPVIVKLYGACLNPPVCLVMEYAEGLSYNNVLHGAEPLPYTAAHAMSACL 135
QY 61 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKSAAW 120
DB 136 QCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKSAAW 195
QY 121 MAPEVFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
DB 196 MAPEVFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPSRPSMEIEIVKIMTHLMRYFPFGADEPLOYEQCQ 228
DB 256 PKPIESLMTRCWSKDPSPSRPSMEIEIVKIMTHLMRYFPFGADEPLOYEQCQ 303

RESULT 6
US-10-106-698-6345
; Sequence 6345, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6345
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (340)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

```

```

US-10-106-698-6345
Query Match      29.6%; Score 371; DB 9; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.4e-27;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVNHNPVIVKLYGACLNPPVCLVMEYAEGLSYNNVLHGAEPLPYTAAHAMS 57
DB 53 EABLISLVSHRNIIQFYGVILEPPNYGIVTEVASYLSGLYDYNINSRSEEM---DMDHMT 109
QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK-- 115
DB 110 WATDVAKGMHYLHMEAPVKVVIHRLDKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLV 167
QY 116 GSAAMAPDEVFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
DB 168 GTFPMMAPEVITQSLPVSSETCDTYSYGVVLEWMLTREVFPKGLG--LQVAVLWVVEKNERL 225
QY 175 PLIKNLPKPIESLMTRCWSKDPSPSRPSMEIEIVKIM 209
DB 226 TIPSSCPRSPFAELLHQCEWADAKKRPSPKQIISIL 316

RESULT 7
US-09-757-982-5
; Sequence 5, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-982-5

Query Match      29.6%; Score 371; DB 9; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.4e-27;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVNHNPVIVKLYGACLNPPVCLVMEYAEGLSYNNVLHGAEPLPYTAAHAMS 57
DB 53 EABLISLVSHRNIIQFYGVILEPPNYGIVTEVASYLSGLYDYNINSRSEEM---DMDHMT 109
QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK-- 115
DB 110 WATDVAKGMHYLHMEAPVKVVIHRLDKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLV 167
QY 116 GSAAMAPDEVFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
DB 168 GTFPMMAPEVITQSLPVSSETCDTYSYGVVLEWMLTREVFPKGLG--LQVAVLWVVEKNERL 225
QY 175 PLIKNLPKPIESLMTRCWSKDPSPSRPSMEIEIVKIM 209
DB 226 TIPSSCPRSPFAELLHQCEWADAKKRPSPKQIISIL 316

RESULT 8
US-10-094-749-2477
; Sequence 2477, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI

```

APPLICANT: SATO, HIROYUKI  
APPLICANT: ISHII, SHIZUKO  
APPLICANT: YAWAMOTO, JUN-ICHI  
APPLICANT: ISONO, YUUKO  
APPLICANT: HIO, YURI  
APPLICANT: OTSUKA, KAORU  
APPLICANT: NAGAI, KEIICHI  
APPLICANT: IRIE, RYOTARO  
APPLICANT: TAMECHIKA, ICHIRO  
APPLICANT: SEKI, NAGHIKO  
APPLICANT: YOSHIKAWA, TSUTOMU  
APPLICANT: OTSUKA, MOTYUKI  
APPLICANT: NAGAHARI, KENJI  
APPLICANT: MASUHO, YASUHIKO  
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
FILE REFERENCE: 084335/0160  
CURRENT APPLICATION NUMBER: US/09/4,749  
CURRENT FILING DATE: 2002-03-12  
PRIOR APPLICATION NUMBER: 60/350,435  
PRIOR FILING DATE: 2002-01-24  
PRIOR APPLICATION NUMBER: JP 2001-328381  
PRIOR FILING DATE: 2001-09-14  
NUMBER OF SEQ ID NOS: 3381  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2477  
LENGTH: 455  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-094-749-2477

Query Match 29.6%; Score 371; DB 12; Length 455;  
Best Local Similarity 37.7%; Pred. No. 1.4e-27;  
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;  
Qy 2 ELRQLSRVNHPIVVKLYGACINP--VCLVMEYABGSGSLYNVLHG--AEPLPYTAHAAMS 57  
Db 53 EAEILSVLSHRNIIQFVGVLPEPNYGVITEYASLSGLYDINSRSEEM--DMDHMT 109  
Qy 58 WLCQSGVAYLHSMQPKALIHRLDKPPNLLVAGGVVVKICDPTGACDIQTHMNNK-- 115  
Db 110 WATDVAKGMYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDPG-ASRPHNHTTHMSLV 167  
Qy 116 GSAAMWAPVEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAPRIMW-AVHNGTRP 174  
Db 168 GTFPWWAPVEIQLPVSGETCDTSYGVVLWEMLTREVFPFKGLEG--LQVAVLVVEKNRL 225  
Qy 175 PLIKNLKPIESLMTRCWSKDPSPQSPSMEIVKIM 209  
Db 226 TIPSSCPSPFAELLHQWEADAKKRPSPFKQIISIL 260

RESULT 9  
US-09-862-027-19  
Sequence 19, Application US/09862027  
Patent No. US20020142428A1  
GENERAL INFORMATION:  
APPLICANT: Hodge, Martin R.  
TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof  
FILE REFERENCE: 35800/234862  
CURRENT APPLICATION NUMBER: US/09/862,027  
CURRENT FILING DATE: 2001-05-21  
PRIOR APPLICATION NUMBER: US 09/345,473  
PRIOR FILING DATE: 1999-06-30  
NUMBER OF SEQ ID NOS: 82  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 19  
LENGTH: 394  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-862-027-19

Query Match 29.0%; Score 363.5; DB 10; Length 394;  
Best Local Similarity 37.7%; Pred. No. 6.5e-27;

Matches 83; Conservative 36; Mismatches 84; Indels 17; Gaps 6;  
Qy 2 ELRQLSRVNHPIVVKLYGACINP--VCLVMEYABGSGSLYNVLHGAEPLPYTAHAAMSWC 59  
Db 50 EAKLFAMLKHPNIIALRGVCLKEPNLCVMEFARGGPNRVLSGRIPPDJ---LVNWA 105  
Qy 60 LQCSQGVAYLHSMQPKALIHRLDKPPNLLV-----AGTVLKIODEGTACD-IQTHM 111  
Db 106 VQIARGNMYLHDEAIVPIIHRDLKSSNIIQKVENGDLNSNKILKITDFGLAREWHRTTK 165  
Qy 112 TNKGSAAWAPVEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAPRIMWAV-HN 170  
Db 166 MSAGTYANWAPVEVIRASMEKSGSDVWSYGVLLWELLTGEVFPFGIDG--LRVAYGVAMN 223  
Qy 171 GTRPPLIKNLKPIESLMTRCWSKDPSPQSPSMEIVKIMT 210  
Db 224 KLALPIPTCTPEPFPAKLMECDWNPDPHSPSFTNILDQLT 263

## RESULT 10

US-10-354-358-24  
Sequence 24, Application US/10354358  
Publication No. US20030157082A1  
GENERAL INFORMATION:  
APPLICANT: Millennium Pharmaceuticals, Inc  
APPLICANT: Hunter, John Joseph  
APPLICANT: MacBeth, Kyle J.  
APPLICANT: Tsai, Fong-Ying  
APPLICANT: Lesoon, Andrea  
APPLICANT: Lightcap, Eric S.  
APPLICANT: Williamson, Mark  
APPLICANT: Rudolph-Owen, Laura A.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING  
TITLE OF INVENTION: CANCER USING 140, 1470, 1686, 2089, 2427, 3702, 5891, 6428,  
TITLE OF INVENTION: 7181, 7660, 25641, 69583, 49863, 8897, 1682, 17667, 9235,  
TITLE OF INVENTION: 703, 14171, 10359, 1660, 1450, 18894, 2088, 32427, 2160,  
TITLE OF INVENTION: 9252, 9389, 1642, 85269, 10297, 1584, 9525, 14124, 4469,  
TITLE OF INVENTION: 8990, 2100, 9288, 64698, 10480, 20893, 33230, 1586, 9943,  
TITLE OF INVENTION: 16334, 68862, 9011, 14031, 6178, 21225, 1420, 32236, 2099,  
TITLE OF INVENTION: 2150, 26583, 2784, 8941, 9811, 27444, 50566 OR 66428 MOLECULES  
FILE REFERENCE: MP102-020PIRNONMIM  
CURRENT APPLICATION NUMBER: US/10/354,358  
CURRENT FILING DATE: 2003-01-30  
PRIOR APPLICATION NUMBER: US 60/353,600  
PRIOR FILING DATE: 2002-01-31  
PRIOR APPLICATION NUMBER: US 60/364,517  
PRIOR FILING DATE: 2002-03-15  
PRIOR APPLICATION NUMBER: US 60/371,075  
PRIOR FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: US 60/371,507  
PRIOR FILING DATE: 2002-04-10  
PRIOR APPLICATION NUMBER: US 60/372,984  
PRIOR FILING DATE: 2002-04-16  
PRIOR APPLICATION NUMBER: US 60/374,194  
PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: US 60/382,995  
PRIOR FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: US 60/385,023  
PRIOR FILING DATE: 2002-05-31  
PRIOR APPLICATION NUMBER: US 60/388,853  
PRIOR FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: US 60/389,395  
PRIOR FILING DATE: 2002-06-17  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 122  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 1036  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-354-358-24

Query Match 28.4%; Score 355; DB 12; Length 1036;



Best Local Similarity 36.1%; Pred. No. 1.4e-25;  
Matches 83; Conservative 40; Mismatches 85; Indels 22; Gaps 6;  
QY 2 ELRQLSRVNHNIIVKLYGACLNPP--VCLVMEYAEAGSGLYVNLHGAEPLPYTYA-----52  
DB 171 EARLFAMLRHPIIIEIRGVCLVLEFARGGALNRALAAANAAPDPRAPGPRAR 230  
QY 53 ---AHAM-SWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAG-----GTVLKICDF 101  
DB 231 RIPPHVLNVNAVQIARGMLYLHEEAFVPIILHRDLKSSNILLLEKIEHDDICNKTUKITDF 290  
QY 102 GTACD-IOTHMTNNKGAAMWAPEFEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGP 160  
DB 291 GLAREWHRTTKMSTAGTYANWAPEVIKSLFSKSGDIWSYGVLLWELLTGEVYPRGIDGL 350  
QY 161 AFRIMWAVHNGTRPPLIKNLKPKPIESLMTRCWSKDPQSRPSMEEIVKIMT 210  
DB 351 AVAYGAVVKNLTL-PIPTSCPEPPFAKLKCKEQQDPHIRPSFALILEQLT 399

## RESULT 11

US-10-014-882-2  
; Sequence 2, Application US/10014882  
; Publication No. US20020107384A1  
; GENERAL INFORMATION:  
; APPLICANT: Hu, Yi  
; APPLICANT: Kleke, James  
; APPLICANT: Donoho, Gregory  
; TITLE OF INVENTION: No. US20020107384A1el Human Kinase and Polynucleotides Encoding b  
; FILE REFERENCE: LEX-0279-USA  
; CURRENT APPLICATION NUMBER: US/10/014,882  
; CURRENT FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: US 60/254,744  
; PRIOR FILING DATE: 2000-12-11  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1036  
; TYPE: PRT  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(1036)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-014-882-2

Query Match 28.4%; Score 355; DB 14; Length 1036;  
Best Local Similarity 36.1%; Pred. No. 1.4e-25;  
Matches 83; Conservative 40; Mismatches 85; Indels 22; Gaps 6;  
QY 2 ELRQLSRVNHNIIVKLYGACLNPP--VCLVMEYAEAGSGLYVNLHGAEPLPYTYA-----52  
DB 171 EARLFAMLRHPIIIEIRGVCLVLEFARGGALNRALAAANAAPDPRAPGPRAR 230  
QY 53 ---AHAM-SWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAG-----GTVLKICDF 101  
DB 231 RIPPHVLNVNAVQIARGMLYLHEEAFVPIILHRDLKSSNILLLEKIEHDDICNKTUKITDF 290  
QY 102 GTACD-IOTHMTNNKGAAMWAPEFEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGP 160  
DB 291 GLAREWHRTTKMSTAGTYANWAPEVIKSLFSKSGDIWSYGVLLWELLTGEVYPRGIDGL 350  
QY 161 AFRIMWAVHNGTRPPLIKNLKPKPIESLMTRCWSKDPQSRPSMEEIVKIMT 210  
DB 351 AVAYGAVVKNLTL-PIPTSCPEPPFAKLKCKEQQDPHIRPSFALILEQLT 399

## RESULT 12

US-10-153-668-436  
; Sequence 436, Application US/10153668  
; Publication No. US20030092616A1  
; GENERAL INFORMATION:  
; APPLICANT: HONDA, Goichi

APPLICANT: MATSUDA, Akio  
APPLICANT: MURAMATSU, Shuji  
APPLICANT: ISHIZAWA, Kenya  
TITLE OF INVENTION: STAT6 Activating Gene  
FILE REFERENCE: 1254-0207P  
CURRENT APPLICATION NUMBER: US/10/153,668  
CURRENT FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: US 60/293,172  
PRIOR FILING DATE: 2001-05-25  
PRIOR APPLICATION NUMBER: US 60/316,031  
PRIOR FILING DATE: 2001-08-31  
PRIOR APPLICATION NUMBER: US 60/328,403  
PRIOR FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: JP 2001-157043  
PRIOR FILING DATE: 2001-05-25  
PRIOR APPLICATION NUMBER: JP 2001-260681  
PRIOR FILING DATE: 2001-08-30  
NUMBER OF SEQ ID NOS: 488  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 436  
LENGTH: 746  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-153-668-436

Query Match 28.2%; Score 353.5; DB 15; Length 746;  
Best Local Similarity 34.7%; Pred. No. 1.3e-25;  
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;  
QY 2 ELRQLSRVNHNIIVKLYGACLNPP--VCLVMEYAEAGSGLYVNLHGAEPLPYTYAAHMSWC 59  
DB 194 DIKHLRLKHPNIIITFKGVCTQAPCYCLIMEFCAQQLYEVLRAGRPV---TPSLLDVMS 250  
QY 60 LQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGVTLKICDFGTACDIQTMTNNK--GS 117  
DB 251 MGIAGGMNLYLHK---IIHRDLKSPN-MLTYDDVVKISDFGSKLSKDSKTKMSFAGT 306  
QY 118 AAWMAPEVFEKSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAPRIMWAV-HNTRPPL 176  
DB 307 VAWMAPEVIRNEPVSEKVDIWSFGVLLWELLTGEIPIYKDVDSA--IIWGVGNSLSLHPV 364  
QY 177 IKNLPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHL 212  
DB 365 PSCPDGFKILLRCQWNSKPRNPSFRQ---ILLHL 397

RESULT 13  
US-10-153-668-324  
; Sequence 324, Application US/10153668  
; Publication No. US20030092616A1  
; GENERAL INFORMATION:  
; APPLICANT: HONDA, Goichi  
; APPLICANT: MATSUDA, Akio  
; APPLICANT: MURAMATSU, Shuji  
; APPLICANT: ISHIZAWA, Kenya  
TITLE OF INVENTION: STAT6 Activating Gene  
FILE REFERENCE: 1254-0207P  
CURRENT APPLICATION NUMBER: US/10/153,668  
CURRENT FILING DATE: 2002-05-24  
PRIOR APPLICATION NUMBER: US 60/293,172  
PRIOR FILING DATE: 2001-05-25  
PRIOR APPLICATION NUMBER: US 60/316,031  
PRIOR FILING DATE: 2001-08-31  
PRIOR APPLICATION NUMBER: US 60/328,403  
PRIOR FILING DATE: 2001-10-12  
PRIOR APPLICATION NUMBER: JP 2001-157043  
PRIOR FILING DATE: 2001-05-25  
PRIOR APPLICATION NUMBER: JP 2001-260681  
PRIOR FILING DATE: 2001-08-30  
NUMBER OF SEQ ID NOS: 488  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 436  
LENGTH: 746  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-153-668-436

```
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 324
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-324

Query Match      28.2%; Score 353.5; DB 15; Length 859;
Best Local Similarity 34.7%; Pred. No. 1.5e-25;
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;

QY      2 ELROLSRVNHPNIVKLYGACLNLP--VCLWMEYABGGSLYNVLHGAELPYYTAHAHMSWC 59
Db      161 DIHLRLKLPNIIITFGVCTQAPCYCILMEFCAGQQLYEVLRAGRPV---TPSLLYDWS 217

QY      60 LQCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTGVLKICDFGTACDIQTHMTNKK--GS 117
Db      218 MGIAGGMNLYLHLK---IIHRLDKSPN-MLITYDDVVVKISDFGTSKELSDKSTKWSFAGT 273

QY      118 AAWMAPEVFGSSNYSEKCDVFSWGIILWEVITRKKPDEIGGPAFRIMWAV-HNGTRPPL 176
Db      274 VAWMAPEVIRNEPVSEKVDIWSFGVLLWELLTGIEIPYKVDSSA--IILHLL 331

QY      177 IKNLKPPIESIMTRCWSKDPSPRSMEEIVKIMTHL 212
Db      332 PSSCPDGFKILLRQCWNKSPNRNRPFRQ---IILHLL 364

RESULT 14
US-10-153-668-438
; Sequence 438, Application US/10153668
; Publication No. US20030092616A1
; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 438
; LENGTH: 892
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-153-668-438

Query Match      28.2%; Score 353.5; DB 15; Length 892;
Best Local Similarity 34.7%; Pred. No. 1.6e-25;
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;

QY      2 ELROLSRVNHPNIVKLYGACLNLP--VCLWMEYABGGSLYNVLHGAELPYYTAHAHMSWC 59
Db      194 DIHLRLKLPNIIITFGVCTQAPCYCILMEFCAGQQLYEVLRAGRPV---TPSLLYDWS 250

QY      60 LQCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTGVLKICDFGTACDIQTHMTNKK--GS 117
Db      251 MGIAGGMNLYLHLK---IIHRLDKSPN-MLITYDDVVVKISDFGTSKELSDKSTKWSFAGT 306

; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 324
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-288-798-324

Query Match      27.9%; Score 349.5; DB 12; Length 1097;
Best Local Similarity 37.4%; Pred. No. 5e-25;
Matches 82; Conservative 35; Mismatches 87; Indels 15; Gaps 5;

QY      2 ELROLSRVNHPNIVKLYGACLNLP--VCLWMEYABGGSLYNVLHGAELPYYTAHAHMSWC 59
Db      191 EAKLFAMLKHPNIIITRGLRGVCLKEPNLCVMEFARGGPNLVLSGRIPDDI---LVNWA 246

QY      60 LQCSQGVAYLHSMOPKALIHRLDKPPNLLV-----AGTVLKI CDFGTACD-IOTHM 111
Db      247 VQIARGMNYLLDEAIVPIIHRDLKSSNIILOKVENGDLNKLKITDFGLAREWHRTTK 306

; NUMBER OF SEQ ID NOS: 498
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 324
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-288-798-324

Query Match      27.9%; Score 349.5; DB 12; Length 1097;
Best Local Similarity 37.4%; Pred. No. 5e-25;
Matches 82; Conservative 35; Mismatches 87; Indels 15; Gaps 5;

QY      2 ELROLSRVNHPNIVKLYGACLNLP--VCLWMEYABGGSLYNVLHGAELPYYTAHAHMSWC 59
Db      191 EAKLFAMLKHPNIIITRGLRGVCLKEPNLCVMEFARGGPNLVLSGRIPDDI---LVNWA 246

QY      60 LQCSQGVAYLHSMOPKALIHRLDKPPNLLV-----AGTVLKI CDFGTACD-IOTHM 111
Db      247 VQIARGMNYLLDEAIVPIIHRDLKSSNIILOKVENGDLNKLKITDFGLAREWHRTTK 306
```



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:28:31 ; Search time 29.2703 Seconds  
(without alignments)  
329.579 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303

Perfect score: 1252  
Sequence: 1 VELQLSRVNHNPVLYGALG.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep:  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep:  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep:  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep:  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep:  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	100.0	579	4	US-09-529-279-4 Sequence 4, Appli
2	1252	100.0	579	4	US-10-158-895-4 Sequence 4, Appli
3	1252	100.0	590	4	US-09-529-279-15 Sequence 15, Appl
4	1252	100.0	590	4	US-10-158-895-15 Sequence 15, Appl
5	371	29.6	455	3	US-09-221-233-5 Sequence 5, Appli
6	371	29.6	455	3	US-09-221-928-5 Sequence 5, Appli
7	371	29.6	455	3	US-09-221-527-5 Sequence 5, Appli
8	371	29.6	455	3	US-09-221-236-5 Sequence 5, Appli
9	371	29.6	455	3	US-09-221-416-5 Sequence 5, Appli
10	371	29.6	455	3	US-09-221-245-5 Sequence 5, Appli
11	371	29.6	455	3	US-09-163-115-5 Sequence 5, Appli
12	371	29.6	455	3	US-09-221-528-5 Sequence 5, Appli
13	371	29.6	455	3	US-09-593-553-5 Sequence 5, Appli
14	371	29.6	455	3	US-09-221-237-5 Sequence 5, Appli
15	371	29.6	455	3	US-09-399-588-2 Sequence 2, Appli
16	363.5	29.0	394	4	US-09-345-473E-19 Sequence 19, Appl
17	353.5	28.2	668	1	US-08-205-018-2 Sequence 2, Appli
18	353.5	28.2	859	1	US-08-395-580-2 Sequence 2, Appli
19	353.5	28.2	859	5	PCT-US95-02792-2 Sequence 2, Appli
20	346	27.6	328	4	US-09-345-473E-18 Sequence 18, Appl
21	338.5	27.0	835	3	US-09-291-839-2 Sequence 2, Appli
22	338.5	27.0	835	4	US-09-458-457-2 Sequence 2, Appli
23	335.5	26.8	835	4	US-09-458-457-8 Sequence 8, Appli
24	323	25.8	821	1	US-07-928-464-2 Sequence 2, Appli
25	323	25.8	821	1	US-08-003-311B-2 Sequence 2, Appli
26	323	25.8	821	1	US-08-261-432-2 Sequence 2, Appli
27	323	25.8	821	5	PCT-US93-07347-2 Sequence 2, Appli

28	320	25.6	263	3	US-09-035-706-5 Sequence 5, Appli
29	320	25.6	263	3	US-08-955-841-5 Sequence 5, Appli
30	320	25.6	263	4	US-09-390-425-5 Sequence 5, Appli
31	320	25.6	263	4	US-09-566-906-5 Sequence 5, Appli
32	309.5	24.7	269	2	US-07-857-224B-79 Sequence 79, Appl
33	307.5	24.6	276	2	US-07-857-224B-72 Sequence 72, Appl
34	306	24.4	275	2	US-07-857-224B-71 Sequence 71, Appl
35	305.5	24.4	590	4	US-09-312-283C-409 Sequence 409, App
36	305.5	24.4	786	4	US-09-503-802-2 Sequence 2, Appli
37	305.5	24.4	787	3	US-09-188-930-334 Sequence 334, App
38	305.5	24.4	787	4	US-09-312-283C-334 Sequence 334, App
39	304.5	24.3	304	2	US-08-701-191A-27 Sequence 27, Appl
40	302.5	24.2	316	1	US-08-278-089A-16 Sequence 16, Appl
41	302.5	24.2	316	2	US-08-838-957A-15 Sequence 15, Appl
42	299.5	23.9	527	4	US-08-428-509A-10 Sequence 10, Appl
43	299.5	23.9	527	4	US-08-232-545-10 Sequence 10, Appl
44	299.5	23.9	527	5	PCT-US95-05008-10 Sequence 10, Appl
45	299.5	23.9	625	1	US-08-391-615-3 Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-09-529-279-4  
; Sequence 4, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1997-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-4

Query Match	100.0%	Score 1252;	DB 4;	Length 579;
Best Local Similarity	100.0%	Pred. No. 4e-133;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	VELQLSRVNHNPVLYGACINPVCLVMEYAEAGSLYNVLHGADPLPYTTAAHMSWCL	60	
Db	76	VELQLSRVNHNPVLYGACINPVCLVMEYAEAGSLYNVLHGADPLPYTTAAHMSWCL	135	
QY	61	QCSQGVAYLHSHQKALIHRLDKPENLLVAGTGVKICDFGTACDIQTHMTNNKGSAAW	120	
Db	136	QCSQGVAYLHSHQKALIHRLDKPENLLVAGTGVKICDFGTACDIQTHMTNNKGSAAW	195	
QY	121	MAPEVFEAGSNVSEKCDVFSWGIILMEVITRRKPPDEIGGPAPRIMWAVHNGTRPPLKNL	180	
Db	196	MAPEVFEAGSNVSEKCDVFSWGIILMEVITRRKPPDEIGGPAPRIMWAVHNGTRPPLKNL	255	
QY	181	PKPIESLMTKWSKDPSPQSPMEETVKIMTHLMRYFPGADEPLQYPCQ	228	
Db	256	PKPIESLMTKWSKDPSPQSPMEETVKIMTHLMRYFPGADEPLQYPCQ	303	

RESULT 2  
US-10-158-895-4  
; Sequence 4, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:

```
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-4

Query Match      100.0%; Score 1252; DB 4; Length 579;
Best Local Similarity 100.0%; Pred. No. 4e-133;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VELRLSRVNHNPNTVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAEPPLPYTAAHANGWCL 60
Db 76 VELRLSRVNHNPNTVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAEPPLPYTAAHANGWCL 135

Qy 61 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTACDIQTHTNNKGSAAW 120
Db 136 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTACDIQTHTNNKGSAAW 195

Qy 121 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL 180
Db 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL 255

Qy 181 PKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLOYPCQ 228
Db 256 PKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLOYPCQ 303

RESULT 4
US-10-158-895-15
; Sequence 15, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-15

Query Match      100.0%; Score 1252; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 4.1e-133;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VELRLSRVNHNPNTVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAEPPLPYTAAHANGWCL 60
Db 76 VELRLSRVNHNPNTVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAEPPLPYTAAHANGWCL 135

Qy 61 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTACDIQTHTNNKGSAAW 120
Db 136 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTACDIQTHTNNKGSAAW 195

Qy 121 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL 180
Db 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL 255

Qy 181 PKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLOYPCQ 228
Db 256 PKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLOYPCQ 303

RESULT 5
US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: WNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
```

```
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-235-5

Query Match      29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVHNHNIVKLYGACLNPP--VCLVMEYAEAGSLYNVLHG--ABPLPYTTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSYDYNINRSEEM---DMDHMT 109
QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLV 167
QY 116 GSAAMMAPEVFEAGSYSEKCDVFSWGIIWVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFFPMAPEVIQSLPVSETCTDYSYGVVLEWMLTREVPFKGLEG--LQVAMLVVEKNERL 225
QY 175 PLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIM 209
Db 226 TIPSSCPRSFAELLHQCEWADAKKRPSPFKQIISIL 260

RESULT 6
US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-928-5

Query Match      29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVHNHNIVKLYGACLNPP--VCLVMEYAEAGSLYNVLHG--ABPLPYTTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSYDYNINRSEEM---DMDHMT 109
QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLV 167
QY 116 GSAAMMAPEVFEAGSYSEKCDVFSWGIIWVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFFPMAPEVIQSLPVSETCTDYSYGVVLEWMLTREVPFKGLEG--LQVAMLVVEKNERL 225
QY 175 PLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIM 209
Db 226 TIPSSCPRSFAELLHQCEWADAKKRPSPFKQIISIL 260

RESULT 7
US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVHNHNIVKLYGACLNPP--VCLVMEYAEAGSLYNVLHG--ABPLPYTTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSYDYNINRSEEM---DMDHMT 109
QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLV 167
QY 116 GSAAMMAPEVFEAGSYSEKCDVFSWGIIWVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFFPMAPEVIQSLPVSETCTDYSYGVVLEWMLTREVPFKGLEG--LQVAMLVVEKNERL 225
QY 175 PLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIM 209
Db 226 TIPSSCPRSFAELLHQCEWADAKKRPSPFKQIISIL 260

RESULT 8
US-09-221-236-5
; Sequence 5, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-236-5

Query Match      29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVHNHNIVKLYGACLNPP--VCLVMEYAEAGSLYNVLHG--ABPLPYTTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSYDYNINRSEEM---DMDHMT 109
QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLV 167
QY 116 GSAAMMAPEVFEAGSYSEKCDVFSWGIIWVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFFPMAPEVIQSLPVSETCTDYSYGVVLEWMLTREVPFKGLEG--LQVAMLVVEKNERL 225
```

```
QY 175 PLIKNLKPPIESLMTRCWSKDPSPORPSMEEIVKIM 209
Db 226 TIPSSCRSPAELLHQWEADAKKRPSPFKQIISIL 260

RESULT 9
US-09-221-416-5
; Sequence 5, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match 29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVNHPIVLYGACINP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPNYGIIVTEYASLSGLYDINSNRSEEM---DMDHMT 109

QY 58 WLCQCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLV 167

QY 116 GSAAWMAPEVFEKGSNYSEKCDVFSWGLLMEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFPMAPEVIQSLPVSETCDTYSYGVVLMELTREVFPFKGLEG--LQVAWLVVVEKNERL 225

QY 175 PLIKNLKPPIESLMTRCWSKDPSPORPSMEEIVKIM 209
Db 226 TIPSSCRSPAELLHQWEADAKKRPSPFKQIISIL 260

RESULT 10
US-09-221-245-5
; Sequence 5, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-245-5

Query Match 29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVNHPIVLYGACINP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPNYGIIVTEYASLSGLYDINSNRSEEM---DMDHMT 109

QY 58 WLCQCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLV 167

QY 116 GSAAWMAPEVFEKGSNYSEKCDVFSWGLLMEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFPMAPEVIQSLPVSETCDTYSYGVVLMELTREVFPFKGLEG--LQVAWLVVVEKNERL 225

QY 175 PLIKNLKPPIESLMTRCWSKDPSPORPSMEEIVKIM 209
Db 226 TIPSSCRSPAELLHQWEADAKKRPSPFKQIISIL 260

RESULT 11
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

Query Match 29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRQLSRVNHPIVLYGACINP--VCLVMEYAEAGGSLYNVLHG--AEPLPYTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPNYGIIVTEYASLSGLYDINSNRSEEM---DMDHMT 109

QY 58 WLCQCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLV 167

QY 116 GSAAWMAPEVFEKGSNYSEKCDVFSWGLLMEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFPMAPEVIQSLPVSETCDTYSYGVVLMELTREVFPFKGLEG--LQVAWLVVVEKNERL 225

QY 175 PLIKNLKPPIESLMTRCWSKDPSPORPSMEEIVKIM 209
Db 226 TIPSSCRSPAELLHQWEADAKKRPSPFKQIISIL 260

RESULT 12
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-528-5
```



```
Query Match          29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELROLRVNHPNIVKLYGACLNPP--VCLVMEYAEAGSLYNNVLHG--AEPLPYTTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYINSRSEEM---DMDHMT 109

QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFNHNHTTHMSLV 167

QY 116 GSAAWMAPEVFECSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFPMAPEVIQSLPVSETCDTYSYGVVLEWMLTREVFPKLEG--LQVAMLVVEKNERL 225

QY 175 PLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIM 209
Db 226 TIPSSCPSPFAELLHQCEWADAKKRPSPKQIISIL 260

RESULT 13
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

Query Match          29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELROLRVNHPNIVKLYGACLNPP--VCLVMEYAEAGSLYNNVLHG--AEPLPYTTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYINSRSEEM---DMDHMT 109

QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFNHNHTTHMSLV 167

QY 116 GSAAWMAPEVFECSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFPMAPEVIQSLPVSETCDTYSYGVVLEWMLTREVFPKLEG--LQVAMLVVEKNERL 225

QY 175 PLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIM 209
Db 226 TIPSSCPSPFAELLHQCEWADAKKRPSPKQIISIL 260

RESULT 14
US-09-221-237-5
; Sequence 5, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
```

```
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-237-5

Query Match          29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELROLRVNHPNIVKLYGACLNPP--VCLVMEYAEAGSLYNNVLHG--AEPLPYTTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYINSRSEEM---DMDHMT 109

QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFNHNHTTHMSLV 167

QY 116 GSAAWMAPEVFECSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFPMAPEVIQSLPVSETCDTYSYGVVLEWMLTREVFPKLEG--LQVAMLVVEKNERL 225

QY 175 PLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIM 209
Db 226 TIPSSCPSPFAELLHQCEWADAKKRPSPKQIISIL 260

RESULT 15
US-09-399-588-2
; Sequence 2, Application US/09399588
; Patent No. 6511825
; GENERAL INFORMATION:
; APPLICANT: Ruggieri, Rosamaria
; APPLICANT: Callow, Marinella
; APPLICANT: Diaz, Paul W.
; TITLE OF INVENTION: No. 6511825el Cell Signaling Polypeptides and Nucleic Acids
; FILE REFERENCE: 1044-US
; CURRENT APPLICATION NUMBER: US/09/399,588
; CURRENT FILING DATE: 1999-09-20
; EARLIER APPLICATION NUMBER: 60/104,088
; EARLIER FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Human J42
US-09-399-588-2

Query Match          29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1.7e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELROLRVNHPNIVKLYGACLNPP--VCLVMEYAEAGSLYNNVLHG--AEPLPYTTAAHAMS 57
Db 53 EAEILSVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYINSRSEEM---DMDHMT 109

QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK-- 115
Db 110 WATDVAKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFNHNHTTHMSLV 167

QY 116 GSAAWMAPEVFECSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 174
Db 168 GTFPMAPEVIQSLPVSETCDTYSYGVVLEWMLTREVFPKLEG--LQVAMLVVEKNERL 225

QY 175 PLIKNLKPPIESLMTRCWSKDPSPQPSMEEIVKIM 209
Db 226 TIPSSCPSPFAELLHQCEWADAKKRPSPKQIISIL 260
```

Thu Dec 4 17:00:22 2003

us-09-830-144-2 copy\_76\_303.ra1

Page 6

Search completed: December 4, 2003, 09:34:53  
Job time : 30.2703 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 11:40:09 ; Search time 459.851 Seconds  
(without alignments)  
1647.862 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHFNIVKLGA.....MTHLMRYFFGADEPQYPCQ 228

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09830144/runat\_03122003\_122511\_23420/app\_query.fasta\_1.654  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODEB=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09830144 -CGN\_1\_1\_428 @runat\_03122003\_122511\_23420  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Length DB ID	Description
1	1252	100.0	1705 14 US-10-283-023-1 Sequence 1, Appli

2	1252	100.0	1705 14 US-10-283-023-3
3	1252	100.0	1788 12 US-10-384-743-14
4	1252	100.0	1788 13 US-10-158-895-14
5	1252	100.0	2656 12 US-10-384-743-3
6	1252	100.0	2656 13 US-10-158-895-3
7	392.5	31.3	1435 12 US-10-210-120-86
8	392.5	31.3	3454 10 US-09-969-347-226
9	392.5	31.3	3454 14 US-10-171-581-312
10	371	29.6	1063 14 US-10-106-698-2068
11	371	29.6	1365 9 US-09-757-982-6
12	371	29.6	2120 9 US-09-757-982-4
13	371	29.6	2251 12 US-10-094-749-838
14	361.5	28.9	3538 12 US-10-288-798-36
15	355	28.4	3111 13 US-10-014-882-1
16	355	28.4	3518 13 US-10-014-882-3
17	355	28.4	5549 12 US-10-354-358-23
18	353.5	28.2	2440 14 US-10-153-668-435
19	353.5	28.2	3365 14 US-10-153-668-323
20	353.5	28.2	3428 14 US-10-153-668-437
21	353.5	28.2	3761 12 US-09-814-353-20138
22	353.5	28.2	6378 12 US-10-252-157-482
23	347	27.7	1428 10 US-09-938-842A-882
24	345.5	27.6	1662 10 US-09-938-842A-1014
25	345.5	27.6	3558 13 US-10-143-133-1
26	338.5	27.0	2505 10 US-09-947-199-3
27	338.5	27.0	3025 10 US-09-947-199-1
28	335.5	26.8	2505 10 US-09-947-199-7
29	335.5	26.8	3026 10 US-09-947-199-9
30	329.5	26.3	1638 10 US-09-938-842A-903
31	328.5	26.2	2211 10 US-09-938-842A-1577
32	323	25.8	2466 12 US-10-171-404A-47
33	313.5	25.0	3286 10 US-09-904-389-1
34	305.5	24.4	1774 11 US-09-866-050A-403
35	305.5	24.4	1774 12 US-10-152-661-403
36	305.5	24.4	2361 12 US-10-128-174-11
37	305.5	24.4	2370 14 US-10-164-080-1
38	305.5	24.4	2370 14 US-10-299-327-1
39	305.5	24.4	3516 11 US-09-866-050A-257
40	305.5	24.4	3516 14 US-10-152-661-257
41	304.5	24.3	2892 10 US-09-938-842A-1073
42	304.5	24.3	3170 12 US-09-971-392-66
43	304.5	24.3	5355 12 US-09-971-392-67
44	304.5	24.3	5615 12 US-10-241-220-18
45	304.5	24.1	3650 12 US-10-366-288-17

## ALIGNMENTS

### RESULT 1

US-10-283-023-1  
; Sequence 1, Application US/10283023  
; Publication No. US20030091573A1  
; GENERAL INFORMATION:  
; APPLICANT: Caroll, Joseph M.  
; TITLE OF INVENTION: Methods and compositions for the  
; diagnosis and treatment of hematological disorders using  
; FILE REFERENCE: MPI01-239P1RM  
; CURRENT FILING DATE: 2002-10-29  
; PRIOR FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1705  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-283-023-1

Alignment Scores: 4.47e-151 Length: 1705  
Pred. No.: 1252.00 Matches: 228  
Score:

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x US-10-283-023-1 (1-1705)

```
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
DB 226 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 285
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
DB 286 TGCTTGATCCAGTGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 345
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
DB 346 CTGCATGTGTGAACCATTCCTATATATCTGCTGCCACGCAATGAGTTGGTGTATA 405
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 406 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 465
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 466 GACCTGAACACCAACCTTACTGCTGTGAGGGGGGACAGTCTCTAAAAATTTGTGAT 525
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
DB 526 TTTGGTACAGCGCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTGG 585
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
DB 586 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATTCAGTGAATAATGTGAGCTTTCAGCTGG 645
QY 141 GlyIleLeuLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
DB 646 GGTATTATCTTTGGGAAGTGATAACGCTCGAAGAACCTTTGATGAGATTGGTGGCCCA 705
QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
DB 706 GCCTTCCGAATCATGTGGCTGTTCAATAATGATGACTCACTTGATGCGGTACTTTCAGAGCAGAT 765
QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerClnArgPro 200
DB 766 CCTAAGCCCATTGAGAGCCTGATGCTGTTGGTCTAAGATCCCTTCCAGCGCCCT 825
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
DB 826 TCAATGGAGGAAATGTGAAAAATAATGACTCACTTGATGCGGTACTTTCAGAGCAGAT 885
```

RESULT 2

US-10-283-023-3

```
; Sequence 3, Application US/10283023
; Publication No. US20030091573A1
; GENERAL INFORMATION:
; APPLICANT: Carroll, Joseph M.
; TITLE OF INVENTION: Methods and compositions for the
; diagnosis and treatment of hematological disorders using
; FILE REFERENCE: MPI01-239FIRM
; CURRENT APPLICATION NUMBER: US/10/283,023
; PRIORITY FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 60/335,044
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1705
```

TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)...(1557)  
US-10-283-023-3

Alignment Scores:

Pred. No.: 4,47e-151 Length: 1705  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x US-10-283-023-3 (1-1705)

```
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
DB 226 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 285
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
DB 286 TGCTTGATCCAGTGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG 345
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
DB 346 CTGCATGTGTGAACCATTCCTATATATCTGCTGCCACGCAATGAGTTGGTGTATA 405
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 406 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 465
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 466 GACCTGAACACCAACCTTACTGCTGTGAGGGGGGACAGTCTCTAAAAATTTGTGAT 525
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
DB 526 TTTGGTACAGCGCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTGG 585
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
DB 586 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATTCAGTGAATAATGTGAGCTTTCAGCTGG 645
QY 141 GlyIleLeuLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
DB 646 GGTATTATCTTTGGGAAGTGATAACGCTCGAAGAACCTTTGATGAGATTGGTGGCCCA 705
QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
DB 706 GCCTTCCGAATCATGTGGCTGTTCAATAATGATGACTCACTTGATGCGGTACTTTCAGAGCAGAT 765
QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerClnArgPro 200
DB 766 CCTAAGCCCATTGAGAGCCTGATGCTGTTGGTCTAAGATCCCTTCCAGCGCCCT 825
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
DB 826 TCAATGGAGGAAATGTGAAAAATAATGACTCACTTGATGCGGTACTTTCAGAGCAGAT 885
QY 221 GluProLeuGlnTyrProCysGln 228
DB 886 GAGCCATTACAGTATCTTGTGAC 909
```

RESULT 3

US-10-384-743-14

```
; Sequence 14, Application US/10384743
; Publication No. US20030162228A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIO
; APPLICANT: TSUCHIYA, MASAYUKI
```

;; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
;; FILE REFERENCE: 053466/0278  
;; CURRENT APPLICATION NUMBER: US/10/384,743  
;; CURRENT FILING DATE: 2003-03-11  
;; PRIOR APPLICATION NUMBER: US/09/529,279  
;; PRIOR FILING DATE: 2000-04-11  
;; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
;; PRIOR FILING DATE: 1998-10-22  
;; PRIOR APPLICATION NUMBER: JP 9/290188  
;; PRIOR FILING DATE: 1997-10-22  
;; NUMBER OF SEQ ID NOS: 48  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 14  
;; LENGTH: 1788  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (7)..(1776)  
US-10-384-743-14  
  
Alignment Scores:  
Pred. No.: 4,78e-151 Length: 1788  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0  
  
US-09-830-144-2\_COPY\_76\_303 (1-228) x US-10-384-743-14 (1-1788)  
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
DB 232 GTAGAGCTTCGGCAGTTATCCGCTGTGAACCATCTAATATTGTAAGCTTTATCGAGCC 291  
  
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
DB 292 TGCITGAATCCAGTGTGCTGTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATGTG 351  
  
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
DB 352 CTGCATGTGTGTGAACCATTCGCAATATTATATCTGCTGCCACGCAATGAGTTGGTGTTA 411  
  
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
DB 412 CAGTGTTCACAGGAGTGGCTTATCTTCACGATGCAACCCAAAGCGCTAATTCACAGG 471  
  
QY 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLysIleCysAsp 100  
DB 472 GACCTGAAACCCACCAACTTATCCGCTGTGAACCATCTAATATTGTAAGCTTTATCGAGCC 291  
  
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120  
DB 532 TTTGTGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTCTCTTGTAT 591  
  
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
DB 592 ATGGCACTGAAAGTTTGAAGGTAGTAAATACAGTGAAATATGACGCTTCACAGTGG 651  
  
QY 141 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160  
DB 652 GGTATATCTTTGGGAAGTGAACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 711  
  
QY 161 AlapheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
DB 712 GCTTTCCGAATCATGTGGCTGTTTATATGTTGTTCTCGACCACTGATATAAAATTTA 771  
  
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200  
DB 772 CCTAAGCCCATTCAGAGCCCTGATGACTCGTTGTTGCTTAAAGATCCCTTCCAGCGCCCT 831  
  
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGluValAsp 220  
DB 944 TTTTGAAGGTAGTAAATTTACAGTGAATAATGTCAGTGAATAATGTCAGTCTCTCAGTGG 651

DB 832 TCAATGGAGGAAATTTGAAAAATATGACTCACTTGTATGCGGTACTTTCCAGGACGAGAT 891  
QY 221 GluProLeuGlnTyrProCysGln 228  
DB 892 GAGCCATTACAGATATCCTTGTGAC 915  
  
RESULT 4  
US-10-158-895-14  
;; Sequence 14, Application US/10158895  
;; Publication No. US20020155624A1  
;; GENERAL INFORMATION:  
;; APPLICANT: ONO, KOICHIRO  
;; APPLICANT: OHTOMO, TOSHIO  
;; APPLICANT: TSUCHIYA, MASAYUKI  
;; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
;; FILE REFERENCE: 053466/0278  
;; CURRENT APPLICATION NUMBER: US/10/158,895  
;; CURRENT FILING DATE: 2002-06-03  
;; PRIOR APPLICATION NUMBER: US/09/529,279  
;; PRIOR FILING DATE: 2000-04-11  
;; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
;; PRIOR FILING DATE: 1998-10-22  
;; PRIOR APPLICATION NUMBER: JP 9/290188  
;; NUMBER OF SEQ ID NOS: 48  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 14  
;; LENGTH: 1788  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (7)..(1776)  
US-10-158-895-14  
  
Alignment Scores:  
Pred. No.: 4,78e-151 Length: 1788  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0  
  
US-09-830-144-2\_COPY\_76\_303 (1-228) x US-10-158-895-14 (1-1788)  
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
DB 232 GTAGAGCTTCGGCAGTTATCCGCTGTGAACCATCTAATATTGTAAGCTTTATCGAGCC 291  
  
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
DB 292 TGCITGAATCCAGTGTGCTGTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATGTG 351  
  
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
DB 352 CTGCATGTGTGTGAACCATTCGCAATATTATATCTGCTGCCACGCAATGAGTTGGTGTTA 411  
  
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
DB 412 CAGTGTTCACAGGAGTGGCTTATCTTCACGATGCAACCCAAAGCGCTAATTCACAGG 471  
  
QY 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLysIleCysAsp 100  
DB 472 GACCTGAAACCCACCAACTTATCCGCTGTGAACCATCTAATATTGTAAGCTTTATCGAGCC 291  
  
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120  
DB 532 TTTGTGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTCTCTTGTAT 591  
  
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
DB 592 ATGGCACTGAAAGTTTGAAGGTAGTAAATACAGTGAAATATGACGCTTCACAGTGG 651  
  
QY 141 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160  
DB 652 GGTATATCTTTGGGAAGTGAACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 711  
  
QY 161 AlapheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
DB 712 GCTTTCCGAATCATGTGGCTGTTTATATGTTGTTCTCGACCACTGATATAAAATTTA 771  
  
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200  
DB 772 CCTAAGCCCATTCAGAGCCCTGATGACTCGTTGTTGCTTAAAGATCCCTTCCAGCGCCCT 831  
  
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGluValAsp 220  
DB 944 TTTTGAAGGTAGTAAATTTACAGTGAATAATGTCAGTGAATAATGTCAGTCTCTCAGTGG 651

```
QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 652 GGTATTATTCTTTGGGAAGTGAATACCGCGTGGAAACCCCTTTGATGAGATTGGTGGCCCA 711
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 712 GCTTTCCGAATCATGTGGCGTGTTCATAATGATGACTGACCAACCACTGATAAAAAATTAA 771
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 772 CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGTCTAAAGATCCTTCCAGCGCCCT 831
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 832 TCAATGAGGAAATGTGAAATAATGACTCACTGTATGCGGTACTTTCAGGAGCAGAT 891
QY 221 GluProLeuGlnTyrProCysGln 228
Db 892 GAGCCATTACAGTATCCTTTGTTCAG 915

RESULT 5
US-10-384-743-3
; Sequence 3, Application US/10384743
; Publication No. US20030162228A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/384,743
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-10-384-743-3

Alignment Scores:
Pred. No.: 8,36e-151 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x US-10-384-743-3 (1-2656)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 408 GTAGAGCTTCGGCAGTTATCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGCC 467
QY 21 CysLeuAnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 468 TCGTTGATCCAGTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATAAATGTG 527
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 528 CTGATGCTGTGTGAACCATTTGCCATATTACTGCTGCCACGCAATGAGTGTGTGTTTA 587
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
```

```
Db 588 CAGTGTTCCTCCAAAGGAGTGGCTTATCTTTCACAGCATGCAACCCAAAGCGCTAATTACACAGG 647
QY 81 AspLeuLysProProAsnLeuLeuValIleGlyGlyThrValLeuLysIleCysAsp 100
Db 648 GACCTGAACACCAACCAACTTACTGCTGGTTGCGAGGGGGACAGTTCTAAAAAATTGTGAT 707
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 708 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACAAAGGGAGTGCTGTGG 767
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db 768 ATGCGACCTCGAAGTTTTTGAAGGTAGTAATTACAGTGAATAAATGTGACGCTTCCAGCTGG 827
QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 828 GGTATTATTCTTTGGGAAGTGAATAACCGCGTGGAAACCCCTTGTGATGAGTTGTGGCCCA 887
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 888 GCITTCGAATCATGTGGCGTGTTCATAATGCTACTCGACCCACCTGATATAAAAAATTAA 947
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 948 CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGTCTAAAGATCCTTCCAGCGCCCT 1007
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 1008 TCAATGAGGAAATGTGAAATAATGACTCACTGTATGCGGTACTTTCAGGAGCAGAT 1067
QY 221 GluProLeuGlnTyrProCysGln 228
Db 1068 GAGCCATTACAGTATCCTTTGTTCAG 1091

RESULT 6
US-10-158-895-3
; Sequence 3, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-10-158-895-3

Alignment Scores:
Pred. No.: 8,36e-151 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x US-10-158-895-3 (1-2656)
```

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
DB GTAGAGCTTGGCAGTTATCCGCTGGAACCATCTATATTTGTAAGCTTTATGAGACC 467  
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40  
DB TGCCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATG 527  
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
DB CTGCATGTGTGTAACCAATGCAATATATCTGCTGCCACGCAATGAGTTGGTGT 587  
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
DB CAGTGTTCCTCCAGAGTGGCTTATCTTACAGCATGCAACCAAGAGCGCTATTTCCAGG 647  
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
DB GACCTGAAACCAACCAAACTTACTGTGTGTGTCAGGGGGACAGTTCTAAAAATTTGTAT 707  
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120  
DB TTTGTATACAGCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTCTGCTTGG 767  
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
DB ATGGCACCTGAAGTTTGTGAGGTAGTAATTTACAGTGAAATGTGACGTCTTACAGCTGG 827  
QY 141 GlyIleIleLeuTrpGluValIleThrArgGlyLysProPheAspGluIleGlyPro 160  
DB GGTATATATCTTTGGGAAGTATACAGCGCTCGAAACCTTTGATGAGATTGGTGGCCCA 887  
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleHisAsnLeu 180  
DB GCTTTCGAATCATGTGGGCTGTTCAATAATGTACTCCACCACTACTATAAAAAATTTA 947  
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200  
DB CCTAAGCCCATGTAGAGCTCATGACTGTGTGTTGGTCTTAAGATCTTCCAGCGCCCT 1007  
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
DB TCAATGAGGAGAAATGTTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGACAGAT 1067  
QY 221 GluProLeuGlnTyrProCysGln 228  
DB GAGCCATTACAGTATCTCTGTGAC 1091

RESULT 7  
US-10-210-120-86  
; Sequence 86, Application US/10210120  
; Publication No. US20030175736A1  
; GENERAL INFORMATION:  
; APPLICANT: Chinnaiyan, Arul M.  
; APPLICANT: Sreekumar, Arun  
; TITLE OF INVENTION: Expression Profile of Prostate Cancer  
; FILE REFERENCE: UM-07221  
; CURRENT APPLICATION NUMBER: US/10/210,120  
; CURRENT FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: US 60/309,581  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: US 60/334,468  
; PRIOR FILING DATE: 2001-11-15  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 86  
; LENGTH: 3435  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-210-120-86  
Alignment Scores:

Pred. No.: 9,11e-40 Length: 3435  
Score: 292.50 Matches: 88  
Percent Similarity: 56.88% Conservative: 35  
Best Local Similarity: 40.55% Mismatches: 77  
Query Match: 31.35% Indels: 17  
DB: 12 Gaps: 6  
US-09-830-144-2\_COPY\_76\_303 (1-228) x US-10-210-120-86 (1-3435)  
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21  
DB 721 GAAGCCCGCTCTTTGGAGCCCTGCGAGCACCCCAACATAATTGCCCTTAGGGCGCCCTG 780  
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39  
DB 781 CTCACCCCCACACCTCTGCTAGTGTGATGATGCGGGGGTGTGCACTGAGCAGG 840  
QY 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59  
DB 841 GTGCTGGCAGGTGCGCGGTGCCACCTCAGTG-----CTGGTCAACTGGGCT 888  
QY 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79  
DB 889 GTGCAGGTGGCGCGGGCATGAACCTACCTACACAATGATGCCCTGTGCCCATCATCCAC 948  
QY 80 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92  
DB 949 CGGACCTCAAGTCCATCAACATCTGATCTCTGGAGGCCATCGAGAACCAACCTCGCA 1008  
QY 93 GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThrHisMet 111  
DB 1009 GACACGGTGTCTCAAGATCACGAGCTTTCGSCCTCGCCCGAGTGCACAAGACCAACAG 1068  
QY 112 ThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyr 131  
DB 1069 ATGAGCGTGGCGGGACCTACGCTGTGATGCGCGGAGGTTATCCGTCTCTCCCTCTTC 1128  
QY 132 SerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArg 151  
DB 1129 TCCAAAAGCAGTGTGTGAGGCTTCGGGTGTCTGTGTGGGAGCTCTCACCGGGGAG 1188  
QY 152 LysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaVal---HisAsn 170  
DB 1189 GTCCCTTACCTGAGATC-----GACGCTTGGCCCGGTGCGGTATGGCGCTATGAT 1242  
QY 171 GlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArg 190  
DB 1243 AAGCTGACGCTGCCATCTCCCTCCACAGTGCCTCCGAGCCCTTTGCCCGCTCTCTGGAG 1302  
QY 191 CysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 207  
DB 1303 TGCTGGGACCCAGACCCCAAGCGGGCGCCAGATTTCGGTATGATCTTTGAAG 1353

RESULT 8  
US-09-969-347-226  
; Sequence 226, Application US/09969347  
; Patent No. US20020115085A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner, Reinhard  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 689290-69  
; CURRENT APPLICATION NUMBER: US/09/969,347  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US/60/237,598  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: US/60/237,604  
; PRIOR FILING DATE: 2000-10-03  
; NUMBER OF SEQ ID NOS: 318  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 226  
; LENGTH: 3454  
; TYPE: DNA

```
; ORGANISM: Homo sapiens
US-09-969-347-226

Alignment Scores:
Pred. No.: 9,18e-40 Length: 3454
Score: 392.50 Matches: 88
Percent Similarity: 56.68% Conservative: 35
Best Local Similarity: 40.55% Mismatches: 77
Query Match: 31.35% Indels: 17
DB: 14 Gaps: 6

US-09-830-144-2_COPY_76_303 (1-228) x US-09-969-347-226 (1-3454)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 721 GAAGCCCGGCTCTTTGGAGCCCTGCAGACCCCAACATAATTCCTTAGGGGCGCTGC 780
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 781 CTCACCCCCACACCTCTGCTAGTAGTGCACCTCAAGTG-----CTGGTCAACTGGCT 840
QY 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
Db 841 GTGCTGCAGGTGCGCGGGTGCACCTCAAGTG-----CTGGTCAACTGGCT 888
QY 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 889 GTGAGGTGCGCGGGCATGAATCTCTACAAATGATGCCCTGTGCCCATCATCCAC 948
QY 80 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92
Db 949 CGGAGCTCAAGTCCATCACTCTGAGGCGCATCGAGACCAACCACTCGCA 1008
QY 93 GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp----IleGlnThrHisMet 111
Db 1009 GACACGGTGTCAAGATCAGGACTTCGGCTCGCGCGGAGTATCCCTCTCTCCCTCTC 1128
QY 112 ThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPheGluGlySerAsnTyr 131
Db 1129 TCACAAAGCAGTGTCTGAGGCTTCGGGTGTCTGTGGAGCTGTACGCGGGGAG 1188
QY 152 LysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaVal---HisAsn 170
Db 1189 GTCCCTTACCGTGAGATC-----GACGCTTGGCGGTGGCGTATGGCTATGAAT 1242
QY 171 GlyThrArgProProLeuLysAsnLeuProLysProLysProLysLeuMetThrArg 190
Db 1243 AAGTGACGTGCGCATTCCTCCACGTGCCCGGAGCCCTTTGCCCGCCCTCTCGAGAA 1302
QY 191 CysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 207
Db 1303 TGCTGGACCCAGACCCCGCGGCGCCAGATTTCGTAGCATCTTGAAG 1353

RESULT 9
US-10-171-581-312
; Sequence 312, Application US/10171581
; Publication No. US2003010426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 312
```

```
; LENGTH: 3454
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: X90846
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-312

Alignment Scores:
Pred. No.: 9,18e-40 Length: 3454
Score: 392.50 Matches: 88
Percent Similarity: 56.68% Conservative: 35
Best Local Similarity: 40.55% Mismatches: 77
Query Match: 31.35% Indels: 17
DB: 14 Gaps: 6

US-09-830-144-2_COPY_76_303 (1-228) x US-10-171-581-312 (1-3454)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 721 GAAGCCCGGCTCTTTGGAGCCCTGCAGACCCCAACATAATTCCTTAGGGGCGCTGC 780
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 781 CTCACCCCCACACCTCTGCTAGTAGTGCACCTCAAGTG-----CTGGTCAACTGGCT 840
QY 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
Db 841 GTGCTGCAGGTGCGCGGGTGCACCTCAAGTG-----CTGGTCAACTGGCT 888
QY 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 889 GTGAGGTGCGCGGGCATGAATCTCTACAAATGATGCCCTGTGCCCATCATCCAC 948
QY 80 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92
Db 949 CGGAGCTCAAGTCCATCACTCTGAGGCGCATCGAGACCAACCACTCGCA 1008
QY 93 GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp----IleGlnThrHisMet 111
Db 1009 GACACGGTGTCAAGATCAGGACTTCGGCTCGCGCGGAGTATCCCTCTCTCCCTCTC 1128
QY 112 ThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPheGluGlySerAsnTyr 131
Db 1069 ATGAGCGCTCGGGGACCTACGCTGATGCGCGGAGTATCCCTCTCTCCCTCTC 1128
QY 132 SerGluLysCysAspValPheSerTrpGlyIleLeuLeuTrpGluValIleThrArgArg 151
Db 1129 TCACAAAGCAGTGTCTGAGGCTTCGGGTGTCTGTGGAGCTGTACGCGGGGAG 1188
QY 152 LysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaVal---HisAsn 170
Db 1189 GTCCCTTACCGTGAGATC-----GACGCTTGGCGGTGGCGTATGGCTATGAAT 1242
QY 171 GlyThrArgProProLeuLysAsnLeuProLysProLysProLysLeuMetThrArg 190
Db 1243 AAGTGACGTGCGCATTCCTCCACGTGCCCGGAGCCCTTTGCCCGCCCTCTCGAGAA 1302
QY 191 CysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 207
Db 1303 TGCTGGACCCAGACCCCGCGGCGCCAGATTTCGTAGCATCTTGAAG 1353

RESULT 10
US-10-106-698-2068/c
; Sequence 2068, Application US/10106698
; Publication No. US2003010960A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
```



```
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MM1-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2068
; LENGTH: 1063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-106-698-2068

Alignment Scores:
Pred. No.: 1,03e-37 Length: 1063
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 14 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-10-106-698-2068 (1-1063)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyAlaCys 21
DB 737 GAGCAGAAATACCTCAGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT 678

QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
DB 677 CTTGAACCTCCCACTAGTGGCATTGTACAGAAATATGCTTCTCGGGATCACTCTATGAT 618

QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
DB 617 TACATTAAACAGTAAACAGAGTGGAGAGATG-----GATATGGATCACTATATGACC 567

QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77
DB 566 TGGCCACTGATGTAGCCAGGAATGCATTATTACATATGGAGGCTCCTGTCAAGGTG 507

QY 78 IleHisArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
DB 506 ATTCACAGAGACCTCAAGTCAAGAAACGTTGTATAGCTGCTGATGGA---GTATTGAAG 450

QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLys-----115
DB 449 ATCTGTGACTTTGGT---GCCTCTCGGTTCATACCAACATCAACACACATGCTCCTTGGTT 393

QY 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135
DB 392 GGAATTTCCCATGGATGGCTCCAGAGTTATCCAGAGTCTCCCTGTGTCAAGAACTTGT 333

QY 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
DB 332 GACACATATTCCTATGTGTGTGTTCTCTGGAGATGCTTAAACAGGAGGTCCTCCCTTTAAA 273

QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
DB 272 GGTITGGAAGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAAGCAGAGATTA 219

QY 175 ProLeuLeuLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
DB 218 ACCATTCAACAGAGTTGCCCCAGAGAGTTTCTGCTGACTGTTTACATCAGTGTGGAGACT 159

QY 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
DB 158 GATGCCAAGAAAGCGCCCATCATTTCAAGCAAAATCATTTCAATCTTG 114

RESULT 11
US-09-757-982-6
; Sequence 6, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
```



Db 646 ATCTGTGACTTTGGT---GCCTCTCGGTTCCTACCAACACACATGCTTGGTT 702  
QY 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135  
Db 703 GGAACCTTTCCCATGATGCTCCAGAGTATCCAGAGTCTCCCTGTGTGTCAGAAACTTGT 762  
QY 136 AspValPheSerTrpGlyLeuLeuTrpGluValIleThrArgArgLysProPheAsp 155  
Db 763 GACACATATCTCTAGTGTGTCTCTGGGAGATGCTAACAGAGGAGGCTCCCTTTAAA 822  
QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174  
Db 823 GGTTTGGAGGA-----TTCAAGTAGCTTGGCTGTAGTGGAAAAAAGAGAGATTA 876  
QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194  
Db 877 ACCATTCAGCAGTGGCCCAAGAGTTTGTCTGGAACCTGTATACATCAGTGTGGGAAGCT 936  
QY 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209  
Db 937 GATGCCAAGAACGCCCATCATTCAGCAAAATCATTTCAATCTGT 981

RESULT 14

US-10-288-798-36  
; Sequence 36, Application US/10288798  
; Publication No. US20030207299A1  
; GENERAL INFORMATION:  
; APPLICANT: BANDMAN, Olga; NGUYEN, Dannel B;  
; APPLICANT: WALIA, Narinder K.; HAFALIA, April J.A.;  
; APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;  
; APPLICANT: GURURAJAN, Rajagopal; DING, Li;  
; APPLICANT: PATTERSON, Chandira; YUE, Henry;  
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.;  
; APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;  
; APPLICANT: LU, Yan; ISON, Craig H.;  
; APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;  
; APPLICANT: AZIMZAI, Valda; BURRILL, John D.;  
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.;  
; APPLICANT: LU, Dyung Aina M.; LAL, Prseti G.;  
; APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;  
; APPLICANT: KEARNEY, Liam; POLICKY, Jennifer L.;  
; APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil  
; TITLE OF INVENTION: HUMAN KINASES  
; CURRENT APPLICATION NUMBER: US/10/288, 798  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: PCT/US01/27219  
; PRIOR FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: US 60/240,542  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: US 60/238,389  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: US 60/236,499  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR APPLICATION NUMBER: US 60/234,902  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: US 60/232,654  
; PRIOR FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: US 60/231,357  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: US 60/229,873  
; PRIOR FILING DATE: 2000-08-31  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PERL Program  
; SEQ ID NO 36  
; LENGTH: 3538  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. US20030207299A1 55052990CB1  
US-10-288-798-36

Alignment Scores:

Pred. No.: 9-48e-36 Length: 3538  
Score: 361.50 Matches: 86  
Percent Similarity: 55.45% Conservative: 36  
Best Local Similarity: 39.09% Mismatches: 82  
Query Match: 28.87% Indels: 17  
DB: 12 Gaps: 5  
US-09-830-144-2\_COPY\_76\_303 (1-228) x US-10-288-798-36 (1-3538)  
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21  
Db 571 GAGGCCAAGCTCTTCGCCATCTGTCGAGCACCACCATCATCTCCCTAAGAGGGGTATGT 630  
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39  
Db 631 CTGAAGAGGAGCCCAACCTCTGCTTGGTCATGGAGTTTGTCTGGTGGAGGACCTTTGAATAGA 690  
QY 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59  
Db 691 GTGTTATCTGGGAAAAGGATTCCTCCCAACATC-----CTGGTGAATTGGGCT 738  
QY 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79  
Db 739 GTGCAGATTGCCAGAGGATGACTTACTTGTATGAGGCAATTGTTCCCATCATCTCAC 798  
QY 80 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92  
Db 799 CGCGACCTTAAGTCAGCAACATATTGATCTCCAGAAAGTGGAGATGGAGACCTGAGC 858  
QY 93 GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThrHisMet 111  
Db 859 AACAGATTCTGAAGATCACTGATTTTGGCTGGCTCGGGAATGGCACCAGACCAAG 918  
QY 112 ThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyr 131  
Db 919 ATGAGTGGCGGAGGACCTGCTGTGATGGACCCAGAGTCATCCGGGCTCCATGTTT 978  
QY 132 SerGluLysCysAspValPheSerTrpGlyLeuIleLeuTrpGluValIleThrArgArg 151  
Db 979 TCCAAAGGCAGTGTGTGGAGCTATGGGTGCTTACTTTGGAGTTGCTGACTGCTGAG 1038  
QY 152 LysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGly 171  
Db 1039 GTGCCCTTTCGAGGCATTGCTGCTTAGCATG-CGC---TTATGGAGTGGCCATGAACAA 1094  
QY 172 ThrArgPro-ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCys 191  
Db 1095 ACTGCGCTTCTCTATCTTCTAGTGGCCAGAACCTTTGCCAACTCATGGAGAGCTG 1154  
QY 191 sTrpSerLysAspProSerGlnArgProSerMetGluGluLeuValLysIleMetThr 210  
Db 1155 CTGGAATCTGATCCCACTCAGCAGCATCTTTCACGAATATCTCGGACCAAGCTTAACC 1212

RESULT 15

US-10-014-882-1  
; Sequence 1, Application US/10014882  
; Publication No. US20020107384A1  
; GENERAL INFORMATION:  
; APPLICANT: Hu, Yi  
; APPLICANT: Kieke, James  
; APPLICANT: Donoho, Gregory  
; TITLE OF INVENTION: No. US20020107384A1  
; FILE REFERENCE: LEX-0279-USA  
; CURRENT APPLICATION NUMBER: US/10/014,882  
; CURRENT FILING DATE: 2001-12-11  
; PRIOR APPLICATION NUMBER: US 60/254,744  
; PRIOR FILING DATE: 2000-12-11  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 3111  
; TYPE: DNA

ORGANISM: homo sapiens  
US-10-014-882-1

Alignment Scores:

Pred. No.: 5,45e-35 Length: 3111  
Score: 355.00 Matches: 83  
Percent Similarity: 53.48% Conservative: 40  
Best Local Similarity: 36.09% Mismatches: 85  
Query Match: 28.35% Indels: 22  
DB: 13 Gaps: 6

US-09-830-144-2\_COPY\_76\_303 (1-228) x US-10-014-882-1 (1-3111)

```

QY      2  GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db      511 GAGGCTCGGCTCTTCGCCATCTGCGCGCACCCCAACATCATCGAGCTGCGCGGTGTGC 570

QY      22  LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db      571 CTGAGGAGCGGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 630

QY      40  ValLeuHisGlyAlaGluProLeuProTyrTyrThrAla----- 52
Db      631 GCGTGGCGCTGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 690

QY      53  -----AlaHisAlaMet---SerTyrCysLeuGlnCysSerGlnGlyValAlaTyr 68
Db      691 CGCATCCCTCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750

QY      69  LeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProProAsnLeuLeu 88
Db      751 CTGATCAGGAGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810

QY      89  LeuValAlaGly-----GlyThrValLeuLysIleCysAspPhe 101
Db      811 CTACTTCAGAGATAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870

QY      102 GlyThrAlaCysAsp---IleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
Db      871 GGGTTGGCGAGGAATGGCACAGACACCAAAATGACACAGCAGCAGCAGCAGCAGCAGCAGC 930

QY      121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
Db      931 ATGCCCGCCGGAAGTATCAAGTCTTCTGTTTCTAAGGGAACGACATCTGGAGCTAT 990

QY      141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db      991 GGAGTGTCTGTGGAACTGCTCACCAGGAGAGTCCCTATCGGGCATTGATGGCTC 1050

QY      161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db      1051 GCCGTGGCTTATGGGGTAGCAGTCAATAAATCACTTTG---CCCATTCATCCACCTGC 1107

QY      181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
Db      1108 CCTGAGCGCTTGGCAAGCTCATGAAGAATGCTGGCAAGACCCCTCATATTCTGTCCA 1167

QY      201 SerMetGluGluIleValLysIleMetThr 210
Db      1168 TCGTTTGCCTTAATCTCGAACAGTTGACT 1197

```

Search completed: December 4, 2003, 14:31:32  
Job time : 467.851 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:28:31 ; Search time 8.72973 Seconds  
(without alignments)  
329.579 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSPFLTQSTWTHPTQSSSSS.....AEFYRLMSVDHGQSVVTP 68

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pap:\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pap:\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pap:\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pap:\*
  - 5: /cgn2\_6/ptodata/1/iaa/PTCUT COMB.pap:\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	504	2	US-08-752-891-2
2	359	100.0	504	2	US-08-752-891-6
3	359	100.0	504	2	US-09-144-178-2
4	359	100.0	504	2	US-09-144-178-6
5	359	100.0	504	3	US-09-406-854-2
6	359	100.0	504	3	US-09-406-854-6
7	359	100.0	504	4	US-09-529-279-2
8	359	100.0	504	4	US-10-158-895-2
9	359	100.0	513	4	US-09-529-279-43
10	359	100.0	513	4	US-10-158-895-43
11	359	100.0	517	4	US-09-529-279-11
12	359	100.0	517	4	US-10-158-895-11
13	77	21.4	16	4	US-09-529-279-41
14	77	21.4	16	4	US-10-158-895-41
15	67.5	18.8	852	2	US-09-070-060-3
16	67.5	18.8	852	3	US-09-357-746-3
17	66	18.4	16	4	US-09-529-279-40
18	66	18.4	16	4	US-10-158-895-40
19	65	18.1	335	4	US-09-252-991A-24494
20	64.5	18.0	244	4	US-09-461-325-463
21	63.5	17.7	240	4	US-09-152-060-100
22	63.5	17.7	335	4	US-09-152-060-64
23	63.5	17.7	391	3	US-08-706-216-6
24	60.5	16.9	854	2	US-09-070-060-4
25	60.5	16.9	854	3	US-09-357-746-4
26	60	16.7	1037	4	US-09-428-711A-21
27	58.5	16.3	1093	3	US-08-545-860D-55

28	58.5	16.3	1093	5	PCT-US94-04496-55	Sequence 55, Appl
29	57.5	16.0	421	4	US-09-322-478-6	Sequence 6, Appli
30	57.5	16.0	701	4	US-09-252-991A-23288	Sequence 23288, A
31	57.5	16.0	856	4	US-09-152-060-77	Sequence 77, Appl
32	56.5	15.7	922	4	US-09-116-473-4	Sequence 4, Appli
33	56.5	15.7	1385	2	US-08-687-399-7	Sequence 7, Appli
34	56	15.6	151	4	US-09-252-991A-31031	Sequence 31031, A
35	56	15.6	181	1	US-08-278-091-15	Sequence 15, Appl
36	56	15.6	181	1	US-08-483-859-15	Sequence 15, Appl
37	56	15.6	181	1	US-08-472-173-15	Sequence 15, Appl
38	56	15.6	181	2	US-08-482-816-15	Sequence 15, Appl
39	56	15.6	181	2	US-08-296-149-15	Sequence 15, Appl
40	56	15.6	181	2	US-08-801-439-15	Sequence 15, Appl
41	56	15.6	181	2	US-08-615-271-15	Sequence 15, Appl
42	56	15.6	181	3	US-09-074-660-15	Sequence 15, Appl
43	56	15.6	181	3	US-09-074-659-15	Sequence 15, Appl
44	56	15.6	181	3	US-09-106-468-15	Sequence 15, Appl
45	56	15.6	181	3	US-09-106-468-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-08-752-891-2  
; Sequence 2, Application US/08752891  
; Patent No. 5837819  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752.891  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-752-891-2

Query Match 100.0%; Score 359; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 2.7e-37;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 60  
Db 437 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 496  
QY 61 EQSVWTAP 68  
Db 497 EQSVWTAP 504

RESULT 2  
US-08-752-891-6  
; Sequence 6, Application US/08752891  
; Patent No. 5837819  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-752-891-6

Query Match 100.0%; Score 359; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 2.7e-37;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 60  
Db 437 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 496  
QY 61 EQSVWTAP 68  
Db 497 EQSVWTAP 504

RESULT 3  
US-09-144-178-2  
; Sequence 2, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-752-891-6

Query Match 100.0%; Score 359; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 2.7e-37;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 60  
Db 437 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 496  
QY 61 EQSVWTAP 68  
Db 497 EQSVWTAP 504

RESULT 4  
US-09-144-178-6  
; Sequence 6, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/144,178  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-144-178-6

Query Match 100.0%; Score 359; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 2.7e-37;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 60  
Db 437 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 496  
QY 61 EQSVWTAP 68  
Db 497 EQSVWTAP 504

RESULT 5  
US-09-144-178-2  
; Sequence 2, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/144,178  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-144-178-2

Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/144,178  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-144-178-2

Query Match 100.0%; Score 359; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 2.7e-37;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 60  
Db 437 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 496  
QY 61 EQSVWTAP 68  
Db 497 EQSVWTAP 504

RESULT 4  
US-09-144-178-6  
; Sequence 6, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/144,178  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-144-178-2

Query Match 100.0%; Score 359; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 2.7e-37;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 60  
Db 437 QSPFTLQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGDGRVPEYVDFAEFYRLWSVDHG 496  
QY 61 EQSVWTAP 68  
Db 497 EQSVWTAP 504

RESULT 5  
US-09-144-178-2  
; Sequence 2, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/144,178  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-144-178-2

Query Match 100.0%; Score 359; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 2.7e-37;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,178
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-144-178-6

Query Match 100.0%; Score 359; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPILTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPILTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEGRVEPYVDFAEFYRLWSVDHG 496

QY 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 5
US-09-406-854-2
; Sequence 2, Application US/09406854
; Patent No. 6140042
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-406-854-2

Query Match 100.0%; Score 359; DB 3; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPILTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPILTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEGRVEPYVDFAEFYRLWSVDHG 496

QY 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 6
US-09-406-854-6
; Sequence 6, Application US/09406854
; Patent No. 6140042
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; CLASSIFICATION:
```

```
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-406-854-6

Query Match      100.0%; Score 359; DB 3; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVEPYVDFAEFYRLWSVDHG 496

QY 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 7
US-09-529-279-2
; Sequence 2, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-529-279-2

Query Match      100.0%; Score 359; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVEPYVDFAEFYRLWSVDHG 496

QY 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 8
US-10-158-895-2
; Sequence 2, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
```

```
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-158-895-2

Query Match      100.0%; Score 359; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.7e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVEPYVDFAEFYRLWSVDHG 496

QY 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 9
US-09-529-279-43
; Sequence 43, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-529-279-43

Query Match      100.0%; Score 359; DB 4; Length 513;
Best Local Similarity 100.0%; Pred. No. 2.8e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVEPYVDFAEFYRLWSVDHG 60
Db 446 QSPFTLTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVEPYVDFAEFYRLWSVDHG 505

QY 61 EQSVVTAP 68
Db 506 EQSVVTAP 513

RESULT 10
US-10-158-895-43
; Sequence 43, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
```



```

; APPLICANT:  TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER:  US/10/158,895
; CURRENT FILING DATE:  2002-06-03
; PRIOR APPLICATION NUMBER:  US/09/529,279
; PRIOR FILING DATE:  2000-04-11
; PRIOR APPLICATION NUMBER:  PCT/JP98/04796
; PRIOR FILING DATE:  1998-10-22
; PRIOR APPLICATION NUMBER:  JP 9/290188
; PRIOR FILING DATE:  1997-10-22
; NUMBER OF SEQ ID NOS:  48
; SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH:  513
; TYPE:  PRT
; ORGANISM:  Homo sapiens
; US-10-158-895-43

```

```

Query Match      100.0%; Score 359; DB 4; Length 513;
Best Local Similarity 100.0%; Pred. No. 2.8e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 QSPPLTLQSTNTHQTSSSSSDGGLFRSRPAHS:PPGEDGRVEPYVDPAFYRLWSVDHG 60
    |||||
Db 446 QSPPLTLQSTNTHQTSSSSSDGGLFRSRPAHS:PPGEDGRVEPYVDPAFYRLWSVDHG 505
    |||||

Qy 61 EQSVVTAP 68
    |||||
Db 506 EQSVVTAP 513
    |||||

```

RESULT 11  
US-09-529-279-11  
; Sequence 11, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 517  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-11

```

Query Match      100.0%; Score 359; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 2.8e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPFLTLQSTNTHQTSSSSDGGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 60
    |||||
Db 437 QSPFLTLQSTNTHQTSSSSDGGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 496
    |||||

Qy 61 EQSVVTAP 68
    |||||
Db 497 EQSVVTAP 504
    |||||

```

RESULT 12  
US-10-158-895-11  
; Sequence 11, Application US/10158895  
; Patent No. 6551840

; GENERAL INFORMATION:  
 ; APPLICANT: ONO, KOICHIRO  
 ; APPLICANT: OHTOMO, TOSHIHIKO  
 ; APPLICANT: TSUCHIYA, MASAYUKI  
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
 ; FILE REFERENCE: 053466/0278  
 ; CURRENT APPLICATION NUMBER: US/10/158,895  
 ; CURRENT FILING DATE: 2002-06-03  
 ; PRIOR APPLICATION NUMBER: US/09/529,279  
 ; PRIOR FILING DATE: 2000-04-11  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
 ; PRIOR FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: JP 9/290188  
 ; PRIOR FILING DATE: 1997-10-22  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 11  
 ; LENGTH: 517  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-158-895-11

```

Query Match      100.0%; Score 359; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 2.8e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 QSPTLTQSTNTHTQSSSSDGLFRSRPAHSLPPGEDGRVEPYVDFAEYRLWSVDHG 60
    |||||
Db 437 QSPTLTQSTNTHTQSSSSDGLFRSRPAHSLPPGEDGRVEPYVDFAEYRLWSVDHG 496
    |||||

Qy 61 EQSVVTAP 68
    |||||
Db 497 EQSVVTAP 504
    |||||

```

RESULT 13

US-09-529-279-41

Sequence 41, Application US/09529279

Patent No. 6451617

GENERAL INFORMATION:

APPLICANT: ONO, KOICHIRO

APPLICANT: OHTOMO, TOSHIHIKO

APPLICANT: TSUCHIYA, MASAYUKI

TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

FILE REFERENCE: 053466/0278

CURRENT APPLICATION NUMBER: US/09/529,279

CURRENT FILING DATE: 2000-04-11

PRIOR APPLICATION NUMBER: PCT/JP98/04796

PRIOR FILING DATE: 1998-10-22

PRIOR APPLICATION NUMBER: JP 9/290188

PRIOR FILING DATE: 1997-10-22

NUMBER OF SEQ ID NOS: 48

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 41

LENGTH: 16

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: peptide

US-09-529-279-41

```

Query Match      21.4%; Score 77; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPFTLQSTNTHTQ 15
    | | | | | | | | | |
Db 2 QSPFTLQSTNTHTQ 16
    | | | | | | | | | |

```

RESULT 14  
US-10-158-895-41

; Sequence 41, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 41  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide  
US-10-158-895-41

Query Match 21.4%; Score 77; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPTLTQSTNTHTQ 15  
Db 2 QSPTLTQSTNTHTQ 16

RESULT 15  
US-09-070-060-3  
; Sequence 3, Application US/09070060  
; Patent No. 5976849  
; GENERAL INFORMATION:  
; APPLICANT: Hustad, Carolyn M.  
; APPLICANT: Ghildyal, Namit  
; TITLE OF INVENTION: Human E3 Ubiquitin Protein  
; TITLE OF INVENTION: Ligase  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.  
; STREET: 1800 Concord Pike  
; CITY: Wilmington  
; STATE: DE  
; COUNTRY: USA  
; ZIP: 19850-5437  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/070,060  
; FILING DATE: 30-APR-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/073,839  
; FILING DATE: 05-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Higgins, Patrick H  
; REGISTRATION NUMBER: 39,709  
; REFERENCE/DOCKET NUMBER: PHM.70312  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302.886.4889  
; TELEFAX: 302.886.8221  
; TELEX:

; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 852 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-09-070-060-3  
  
Query Match 18.8%; Score 67.5; DB 2; Length 852;  
Best Local Similarity 31.1%; Pred. No. 5.1;  
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;  
  
QY 2 SPTLTQSTNTHTQSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48  
Db 232 SSTGSLPPPTNTNTSEGATSLIPLITISGGSGRPLNPVTOAPLPPGWEQRVD----- 286  
  
QY 49 AEFYRLMSVDHGQ 62  
Db 287 -QHGRVYVDHVEK 299

Search completed: December 4, 2003, 09:34:53  
Job time : 8.72973 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 04:21:38 ; Search time 2450.23 Seconds  
(without alignments)  
6784.777 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091

Perfect score: 684  
Sequence: 1 gtgagcttcggcaggtatc.....cattacagtatcctgtcag 684

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22791392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_man:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	622	90.9	1201	13 BX364931	BX364931 BX364931
2	586.4	85.7	2994	11 AK053911	AK053911 Mus muscu
3	583.2	85.3	2976	11 AK046565	AK046565 Mus muscu
4	581.6	85.0	1221	13 B0511353	B0511353 AGENCOURT

5	543.6	79.5	739	12	BI696710
6	543	79.4	1062	12	BM554120
7	530	77.5	600	14	CB112561
8	509.8	74.5	1087	9	AL525728
9	482.6	70.6	604	12	BM426610
10	472.6	69.1	542	10	AW960377
11	468.2	68.5	1118	9	AL550589
12	448.2	65.5	839	13	BU418819
13	446.4	65.3	910	10	BF780358
14	444.6	65.0	750	13	BU104161
15	433.6	63.4	590	14	CB583299
16	408	59.6	646	12	B073883
17	407	59.5	686	10	BG548917
18	406.4	59.4	688	12	B0762988
19	406.4	59.4	696	12	B074867
20	405.4	59.3	1006	13	BQ219348
21	399	58.3	771	12	BI093821
22	390.4	57.1	1016	9	AL520975
23	374.2	54.7	742	13	BU058580
24	372	54.4	668	13	BU259310
25	370.4	54.2	1157	14	CD508913
26	343.6	50.2	641	13	BQ386875
27	327.2	47.8	1054	13	BQ067693
28	317	46.3	747	12	BI334774
29	309	45.2	760	12	BG773488
30	300.4	43.9	710	13	EX473160
31	298	43.6	711	9	AV718168
32	296.8	43.4	586	12	BU035129
33	296.4	43.3	924	14	CA984799
34	290.8	42.5	745	12	BI767913
35	287.8	42.1	822	12	BI757369
36	286	41.8	560	2	HS0084713
37	269.2	39.4	820	14	CB995067
38	266.8	39.0	858	12	BG916821
39	264.6	38.7	1033	10	BE882182
40	258.6	37.8	588	13	BU381175
41	257.8	37.7	538	13	BQ557823
42	257	37.6	894	14	CA792459
43	254.6	37.2	788	14	CD350811
44	251.8	36.8	1191	12	BM472107
45	247.8	36.2	707	13	BW233644

## ALIGNMENTS

RESULT 1  
BX364931  
LOCUS  
DEFINITION  
BX364931 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI032YB01 5-PRIME, mRNA sequence.  
1201 bp mRNA linear EST 05-MAY-2003

ACCESSION  
BX364931  
VERSION  
BX364931.1  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens

REFERENCE  
1 (bases 1 to 1201)

AUTHORS  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE  
Full-length cDNA libraries and normalization

JOURNAL  
Unpublished

COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequenage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3613.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS1A1008ZH01QPI&cluster=3613.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Paraday Avenue Genoscope sequence ID : CSIAI008ZH01QP1.

```
FEATURES
Source
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI032YB01"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      290 a   253 c   320 g   299 t   39 others
ORIGIN
Query Match      90.9%; Score 622; DB 13; Length 1201;
Best Local Similarity 96.9%; Pred. No. 1.8e-175;
Matches 663; Conservative 11; Mismatches 6; Indels 4; Gaps 4;
QY      1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTTCTAAAGCTTTATGAGCC 60
Db
446 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTTCTAAAGCTTTATGAGCC 505
QY      61 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATG 120
Db
506 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATG 565
QY      121 CTGATGTGTGTAACCATTCATATATATCTGTGCGCCAGCAATGATTTGGTGT 180
Db
566 CTGCATGTGTGTAACCATTCATATATATCTGTGCGCCAGCAATGATTTGGTGT 625
QY      181 CAGTGTTCCTCAAGGAGTGCTTATCTTCACATGTCACCAACCAAGCGTAATTCACAG 240
Db
626 CAGTGTTCCTCAAGGAGTGCTTATCTTCACATGTCACCAACCAAGCGTAATTCACAG 685
QY      241 GACCTGAACACCAACCTTACTGCTGTGTGAGGGGGACAGTCTCTAAATTTGTGAT 300
Db
686 GACCTGAACACCAACCTTACTGCTGTGTGAGGGGGACAGTCTCTAAATTTGTGAT 745
QY      301 TTGTGTACAGCTGTGACATTCAGACACATGACCAATPAACAAGGGAGTCTGCTTGG 360
Db
746 TTGTGTACAGCTGTGACATTCAGACACATGACCAATPAACAAGGGAGTCTGCTTGG 805
QY      361 ATGGCACCTGAAGTTTGAAGGTACTATTAACGTGAAAATGTCAGCTCTTCAGCTGG 420
Db
806 ATGGCACCTGAAGTTTGAAGGTACTATTAACGTGAAAATGTCAGCTCTTCAGCTGG 865
QY      421 GGTATTATTCTTTGGGAAGTGATAACGGCTCGGAACCCCTTTTGATGAGATTGGTGGCCCA 480
Db
866 GGTATTATTCTTTGGGAAGTGATAACGGCTCGGAACCCCTTTTGATGAGATTGGTGGCCCA 925
QY      481 GCTTTCCGAATCATGTGGCTGTTTCATATGTGATCTGACACCACTGATATAAAATTTA 540
Db
926 GCTTTCCGAATCATGTGGCTGTTTCATATGTGATCTGACACCACTGATATAAAATTTA 985
QY      541 CCTAAGCCCATTCAGAGCTGATGACTCGTGTGTGCTTAAGATCCCTCCAGCGCCCT 600
Db
986 CCTAAGCCCATTCAGAGCTGATGACTCGTGTGTGCTTAAGATCCCTCCAGCGG-CCY 1044
QY      601 TCAATGGAGGAATTTGAAATAATGACTCATTGATCGCGTACTTTCCAGGAGCAGAT 660
Db
1045 TCAATGGAGGAATTTGAAATAATG-MTCMTTGTATGGCGGACTTTCMGGAG-AGAT 1102
QY      661 GAGCCATACAGTATCCTTGTCAG 684
Db
1103 GAGCC-WTAMAGTWTCTTKTRAG 1125
```

RESULT 2  
AK053911  
LOCUS AK053911 2994 bp mRNA linear HTC 05-DEC-2002  
DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched

library, clone:El30320K22 product:mitogen activated protein kinase kinase kinase 7, full insert sequence.

```
ACCESSION AK053911
VERSION AK053911.1 GI:26343862
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S., Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R., Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Kuenli,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M., Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H., Baldairelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N., Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seva,T., Shibata,Y., Storch,K.P., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Wittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kotsuki,S. and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
PUBMED 121217851
```

Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saito, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Takagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

## FEATURES

Location/Qualifiers

1..2994

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="FANTOM DB:EI130320K22"

/db\_xref="taxon:10090"

/clone="EI130320K22"

/tissue\_type="eyeball"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="0 day neonate"

143..1882

/notes="unnamed protein product; mitogen activated protein kinase kinase kinase 7 (MGP) [MGI:1346877, GB|D76446, evidence: BLASTN, 99%, match=2403] putative"

/codon\_start=1

/protein\_id="BAC35588.1"

/db\_xref="GI:26343863"

/translation="MSTASASSSSSSEMIEAPSQVLFNFEIDYKEIEVEWGR  
GAFGVCKAKRKAQVAIKQIESERKATIVELRQLRVNHNIVKLYGACLNPLCL  
VMEYAGSLYLVNHAEPPIYPTAAHMSWCLQSQGVAYLHMQPKALHILDKLP  
NLLVAGTGLVKICDFGTACDIQTHMTNNGSAWMAPEVFGNSYSEKCDVFWGII  
LWEVITRKEFDIGFAPRIMVAVHNGTPPLIKNPKETESIMTLCWGSQDSQRS  
MEIVTLMHMYFPCADEPLOVPCOYSDGQNSATSTGSMFIDIASTNTSKSDTN  
MEQVPATNDITKLESLLKNOAKQOESGRLSLGASRGSVESLPTSTGKMSADM  
SEIARIVATAGNGQRRRSIQDLTVTGEPGVSSSSSPSVRMITTSPTSEKPAR  
SHPTWPDSDTIDTSDNSIMAYLTLQLOPLAPCNKESMAVFOHCKMAQEIYMK  
VQTEIALLLQKQELVAELDQEDQKQNTSLRQVHEKLLDENKSLSTYYQQCKKQLE  
VIRSQQKQKQTS"

BASE COUNT 854 a 657 c 749 g 734 t

## ORIGIN

Query Match 85.7%; Score 586.4; DB 11; Length 2994;  
Best Local Similarity 91.1%; Pred. No. 1.7e-164;  
Matches 623; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 1 GTGAGCTTCGGCAGTTATCCGCTGTGAACCATCTTATATTTGTAAGCTTTATGGAGCC 60  
DB 368 GTGGAGCTCCGGCAGTTGTCACGTGTGAACCATCTTATATTTGTAAGCTTTATGGAGCC 427

QY 61 TGCCTTGAATCCAGTGTGCTTGTGATGAATATGCTGAAGGGGGCTCTTTATATAATGTG 120  
DB 428 TGCCTGATCCAGTATGCTTGTGATGAATATGATGAGAGGGGGCTCTTATATAATGTG 487

QY 121 CTGATGCTGTGTAACCATATGTCATATATATGCTGCGCCACGCAATGAGTGTGTTTGA 180  
DB 488 CTGATGCTGTGTAACCATATGCTTACTACATGCTGCTCATGCCATGAGTGTGTTTGA 547

QY 181 CAGTGTTCACAGGAGTGGCTTATCTTACAGCATGCAACCAAGCGCTAATTCACAGG 240

Db 548 CAGTGTTCACAGGAGTGGCTTACCTGACACATGCAAGCCCAAGAGCGTGAATTCACAGG 607

QY 241 GACCTGAACACCAACCAACTTACTGCTGTGAGGGGGGACAGTCTTAAATAATTTGTGAT 300

Db 608 GACCTCAAGCCTCCAACTTGTGCTGTGAGGGGGGACAGTCTTAAATAATTCGCGAT 667

QY 301 TTGTGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTGCTTGG 360

Db 668 TTGTGTACAGCCTGTGACATCCAAACACACATGACCAATAACAAGGGGAGTGTGCTTGG 727

QY 361 ATGGCACCCTGAAGTTTGAAGTAGTAATTAACAGTGAATAATGTCAGTCTTTCAGTGG 420

Db 728 ATGGCCTCTGAAGTATTTGAAGTAGTAATTAACAGTGAATAATGTCAGTCTTTCAGTGG 787

QY 421 GGTATTTATTTCTTGGGAAGTGAATGACCGTTCGAAACCTTTGATGAGTGTGTCGCCCA 480

Db 788 GGTATTTATCTCTGGGAAGTGAATGACCGCGGAAACCTTTGATGAGTGTGTCGCCCA 847

QY 481 GCTTTCCGAATCATGTGGGCTGTTCATATGTTGCTGACCCACCTGATATAAAATTTA 540

Db 848 GCTTTGAGATCATGTGGGCTGTTCATATGTTGCTGACCCACCTGATCAAAATTTA 907

QY 541 CCTAAGCCCATGAGAGCCTGATGCTGTGTTGTTGTTAAAGATCTTCCCGAGCGCCT 600

Db 908 CCTAAGCCCATGAGAGCTTGTGACACGCTGTGTTGTTAAAGACCATCTCAGCGCCT 967

QY 601 TCAATGAGGAAATTTGTAATAATGACTCACTGATCGGTGCTTCCAGGAGCAGAT 660

Db 968 TCAATGAGGAAATTTGTAATAATGACTCACTGATCGGTGCTTCCAGGAGCAGAT 1027

QY 661 GAGCCATTACAGTATCTTCTGTCAG 684

Db 1028 GAGCCATTACAGTATCTTCTGTCAG 1051

## RESULT 3

AK046565

LOCUS

DEFINITION

Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length

enriched library, clone:B43010B05 product:mitogen activated

protein kinase kinase 7, full insert sequence.

ACCSSION AK046565.1 GI:26091589

VERSION HTK; CAP trapper.

KEYWORDS Mus musculus (house mouse)

SOURCE Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

PUBMED 9279253

REFERENCE 2

10349636

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL 20499374

MEDLINE 11042159

PUBMED 11042159

REFERENCE 3

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,

Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multipillar sequencer

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

20530913  
11076861  
4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Aehburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hune, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, J., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S., and Hayashizaki, Y.

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6921), 685-690 (2001)  
21085660  
11217851

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 2976)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE  
JOURNAL

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. .2976  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="FANTOM DB:B430101B05"  
/db\_xref="taxon:10090"  
/clone="B430101B05"  
/sex="male"  
/tissue\_type="adipose"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="4 days neonate"

FEATURES  
source

misc\_feature 124..1862  
/note="mitogen activated protein kinase 7  
(MGD|MG1:1346877, GB|D76446, evidence: BLASTN, 99%,  
match=2403)  
putative"  
polyA\_signal 2963..2968  
/note="putative"  
polyA\_site 2976  
/note="putative"  
BASE COUNT 850 a 651 c 743 g 732 t  
ORIGIN  
Query Match 85.3%; Score 583.2; DB 11; Length 2976;  
Best Local Similarity 90.8%; Pred. No. 1.5e-163;  
Matches 621; Conservative 0; Mismatches 63; Indels 0; Gaps 0;  
QY 1 GTAGAGCTTCGGCAGTATATCCCGTGTGAACCATCTTAATATCTAAAGCTTTATGAGCC 60  
Db 348 GTGAGCTCCGGAGTGTGCACGTGTGAACCATCTTAATATCTAAAGCTTTATGAGCC 407  
QY 61 TGTTCGAATCCAGTGTGCTTGTGTGAATATGTCTGAAGGGGGCTCTTTATATAATGTG 120  
Db 408 TGCCTGAATCCAGTATGTCTTGTGTGGATATGACAGGGGGCTCAITGTATATGTG 467  
QY 121 CTGCATGCTGCTGAACCATTCGCAATATATCTGCTGCCACCCCAATGATGCTGGTGT 180  
Db 468 CTGCATGCTGCTGAACCATTCGCTTACTACCTGCTGCTCATGCCATGAGCTGGTGT 527  
QY 181 CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCCCAAGCGCTTAATTCACAGG 240  
Db 528 CAGTGTTCCTCAAGGAGTGGCTTACCTGACAGCATGCGAGGGGCTGATTCACAGG 587  
QY 241 GACCTGAAACCCAACTTACCTGCTGTGTCAGGGGGACAGTCTTCAAAATTTGTGAT 300  
Db 588 GACCTCAAGCTCCAACTTGTCTGCTGTGTCAGGAGGACAGTCTTCAAAATCTCGAT 647  
QY 301 TTGTGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGTGG 360  
Db 648 TTGTGTACAGCTGTGACATTCACATCCAAACACACATGACCAATTAACAGGGGAGTGTCTGTGG 707  
QY 361 ATGGCACCTTGAAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 420  
Db 708 ATGGCGCATGAAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 767  
QY 421 GGTATATTTCTTTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 480  
Db 768 GGTATATTTCTTCTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 827  
QY 481 GCTTTCCGATCATGTGGCTGTTCATATGTGTACTCGACCACTCTGATATAAAATTTA 540  
Db 828 GCTTTTCAGATCATGTGGCTGTTCATATGTGTACTCGACCACTCTGATATAAAATTTA 887  
QY 541 CCTAAGCCCATGAGAGCTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 600  
Db 888 CCTAAGCCCATGAGAGCTGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 947  
QY 601 TCAATGGAGGAATTTGTGAAAATATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 660  
Db 948 TCAATGGAGGAATTTGTGAAAATATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 1007  
QY 661 GAGCCATACAGTATCCTTGTGAC 684  
Db 1008 GAGCCGTACAGTATCCTTGTGAC 1031  
RESULT 4  
BUS11353  
LOCUS BUS11353  
DEFINITION AGNCOURT\_10128981 NIH MGC\_134 Mus musculus cDNA clone  
IMAGE:6506145 5', mRNA sequence.  
ACCESSION BUS11353  
VERSION BUS11353.1 GI:22817586  
KEYWORDS EST.



```

Db 121 TGCTTACTACACTGCTGCTCATGCCATGAGTGGTGTACAGTGTCCAGGAGTAGTG 180
QY 199 GCTTATCTTCACAGCATGACACCAAGAGCCTAATTCACAGGAGCCTGAAACCAAC 258
Db 181 GCTTACTGTCACAGCATGACAGCCCAAGAGCCTGATTCACAGGAGCCTCAAGCCTCAAC 240
QY 259 TTACTGCTGTTGACGGGGGACAGTCTTAAATTTGATTTGGTACAGCCTGTGAC 318
Db 241 TTGCTGCTGTTGACAGGAGGACAGTCTTAAATTTGATTTGGTACAGCCTGTGAC 300
QY 319 ATTACAGACACATGACCAATTAACAGGGGAGTGGCTGTGGATGGCACCTGAAGTTTTT 378
Db 301 ATCCAAACACACATGACCAATTAACAGGGAGTGGCTGTGGATGGCGCTGAAGTTTTT 360
QY 379 GAAGGTAGTAATTAACAGTGAATAATGAGCTTTCAGTGGGGTATTAATTTGGGAA 438
Db 361 GAAGGTAGTAATTAACAGTGAATAATGAGCTTTCAGTGGGGTATTAATTTGGGAA 420
QY 439 GTGATAACCGCTCGGAAACCCCTTTGATGAGATTGGTGGCCAGCTTTCGGAATCATGTGG 498
Db 421 GTGATAACCGCTCGGAAACCCCTTCGATGAGATCGTGGCCAGCTTTCGGAATCATGTGG 480
QY 499 GCTGTTCAATAGTGTACTCGACCACTGATCAAAAAATTTTACCTAAGCCCATTTGAGAGC 558
Db 481 GCTGTTCAATAGTGTACTCGACCACTGATCAAAAAATTTTACCTAAGCCCATTTGAGAGC 540
QY 559 CTGATGACTCGTGTGTTGTTCTAAAGATCTTCCAGCGGCCCTCAATGAGGAAATTTGG 618
Db 541 TTGATGACAGCTGTGTGTTCTAAGGACCCATCTCAGGGCCCTTCAATGAGGAAATTTGG 600
QY 619 AAAATAATGACTCACCTTGTGATGCGGTACTTTCAGGAGCAGATGAGCCATTCAGTATCCT 678
Db 601 AAAATAATGACTCACCTTGTGATGCGGTACTTTCAGGAGCAGATGAGCCATTCAGTATCCT 660
QY 679 TGTCAG 684
Db 661 TGTCAG 666

```

```

RESULT 6
BM554120
LOCUS
DEFINITION
AGENCOURT 6546903 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742353
5', mRNA sequence.
ACCESSION
BM554120
VERSION
BM554120.1 GI:18793446
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1062)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILLNL at:
http://image.llnl.gov
Plate: L1AM12760 row: f column: 18
High quality sequence stop: 730.
Location/Qualifiers
1. 1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5742353"

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

FEATURES
source

```

```

/tissue_type="medulla"
/lab_host="DH10B"
/clone_lib="NIH MGC 119"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH MGC Library."
BASE COUNT 253 a 243 c 298 g 266 t
ORIGIN

```

```

Query Match 79.4%; Score 543; DB 12; Length 1062;
Best Local Similarity 96.3%; Pred. No. 9.7e-152;
Matches 619; Conservative 0; Mismatches 17; Indels 7; Gaps 6;

QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCC 60
Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCC 467
QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120
Db 468 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 527
QY 121 CTGCATGTGTGTAACCATTTGCCATATTACTGTGCCCCACGCAATGAGTTGGTGTTA 180
Db 528 CTGCATGTGTGTAACCATTTGCCATATTACTGTGCCCCACGCAATGAGTTGGTGTTA 587
QY 181 CAGTGTTCACAGGAGTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAGG 240
Db 588 CAGTGTTCACAGGAGTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAGG 647
QY 241 GACCTGAAACACCAACAACTTACTGTGTGTCAGGGGGGACAGTCTTAAATAATTTGTGAT 300
Db 648 GACCTGAAACACCAACAACTTACTGTGTGTCAGGGGGGACAGTCTTAAATAATTTGTGAT 707
QY 301 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAGGGGAGTGTCTGTGG 360
Db 708 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAGGGGAGTGTCTGTGG 767
QY 361 ATGGCACCTGAGTTTGAAGGTAGTAAATACAGTGAATAATGACGTCTTCAGCTGG 420
Db 768 ATGGCACCTGAGTTTGAAGGTAGTAAATACAGTGAATAATGACGTCTTCAGCTGG 827
QY 421 GGTATTATTCTTTGGGAAGTGATAACGGTCGG-AAAACCTTTGATGAGATTGGTGG-CC 478
Db 828 GGTATTATTCTTTGGGAAGTGATAACGGTCGG-AAAACCTTTGATGAGATTGGTGG-CC 887
QY 479 CAGCTTTCCGAATCATGTGGGCTGTTCATAATGTTGACTGCACCACTG-ATAAAAAAT 537
Db 888 CAGCTTTCCGAATCATGTGGGCTGTTCATAATGTTGACTGCACCACTG-ATAAAAAAT 947
QY 538 TTACCTAAGCCCATTTG-AGAGCCTCATGCTGTTT--GGTCTAAGATCCTTCCAG 594
Db 948 TTACCTAAGCCCATTTG-AGAGCCTCATGCTGTTT--GGTCTAAGATCCTTCCAG 1007
QY 595 CGCCCTTCAATGGA-GGAAATTTGAAAAATTAATGACTCACTTG 636
Db 1008 CGCCCTTCAATGGA-GGAAATTTGAAAAATTAATGACTCACTTG 1050

```

```

RESULT 7
CB112561
LOCUS
DEFINITION
K-EST0154393 L6ChocK0 Homo sapiens cDNA clone L6ChocK0-4-G09 5',
mRNA sequence.
ACCESSION
CB112561
VERSION
CB112561.1 GI:27938368
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

```



ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 600)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 4 row: G column: 09  
High quality sequence stop: 600.  
Location/Qualifiers  
1. .600  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L6ChoCKO-4-G09"  
/sex="M"  
/cell\_line="Cho-CK"  
/lab\_host="Top10F"  
/clone\_lib="L6ChoCKO"  
/notes="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI;  
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including  
EcoRI site by treatment of T4 RNA ligase and the first  
strand cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transposition of  
competent cells E. coli Top10F' by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."  
BASE COUNT 164 a 113 c 154 g 169 t  
ORIGIN  
Query Match 77.5%; Score 530; DB 14; Length 600;  
Best Local Similarity 100.0%; Pred. No. 5.5e-148;  
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGCC 60  
DB 71 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGCC 130  
QY 61 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
DB 131 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 190  
QY 121 CTGATCGTGTGAACATTCGCAATATATCTCTGCCACGCAATGCTGTGTTA 180  
DB 191 CTGATCGTGTGAACATTCGCAATATATCTCTGCCACGCAATGCTGTGTTA 250  
QY 181 CAGTGTTCACAGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAAATTCACAGG 240  
DB 251 CAGTGTTCACAGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAAATTCACAGG 310  
QY 241 GACCTGAAACCAACCAACTTACTGTGTGTGACGGGGGACAGTTCTAAAAATTGTGAT 300  
DB 311 GACCTGAAACCAACCAACTTACTGTGTGTGACGGGGGACAGTTCTAAAAATTGTGAT 370  
QY 301 TTGTGTACAGCTGTGACATTCACACACATGACCATTAACACAGGGAGTGTCTGTGG 360

DB 371 TTGTGTACAGCTGTGACATTCAGACACACATGACCAATAACAGAGGGAGTCTCTTGG 430  
QY 361 ATGGCAGCTGAGTGTGTTTGAAGGTAGTAAATACAGTGAATAATGACGCTTTCAGCTGG 420  
DB 431 ATGGCAGCTGAGTGTGTTTGAAGGTAGTAAATACAGTGAATAATGACGCTTTCAGCTGG 490  
QY 421 GGTATTATTCTTTGGGAAGTGATACAGGCTCGGAAACCCCTTTGATGAGATTGGTGGCCCA 480  
DB 491 GGTATTATTCTTTGGGAAGTGATACAGGCTCGGAAACCCCTTTGATGAGATTGGTGGCCCA 550  
QY 481 GCTTTCGAATCATGTGGGCTTTCATAATGATGCTACGACCACTGAT 530  
DB 551 GCTTTCGAATCATGTGGGCTTTCATAATGATGCTACGACCACTGAT 600  
RESULT 8  
AL525728 1087 bp mRNA linear EST 23-MAY-2003  
LOCUS AL525728 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
DEFINITION cDNA clone CSODC013YE20 5-PRIME, mRNA sequence.  
ACCESSION AL525728 GI:31063592  
VERSION AL525728.2  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1087)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT On Feb 13, 2001 this sequence version replaced gi:12789221.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3613.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSODC013BC10QPl&cluster=3613.r. Contact :  
Peng Liang Email : filiang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CSODC013BC10QPl.  
FEATURES  
Location/Qualifiers  
1. .1087  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSODC013YE20"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/notes="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 267 a 247 c 289 g 265 t 19 others  
ORIGIN  
Query Match 74.5%; Score 509.8; DB 9; Length 1087;  
Best Local Similarity 98.1%; Pred. No. 9.5e-142;  
Matches 574; Conservative 4; Mismatches 1; Indels 6; Gaps 6;  
QY 1 GTAGAGCTTCGGCAG-TTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGC 59  
DB 489 GTAGAGCTTCGGCAGATTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGC 548  
QY 60 CTGCTTGAATCCAGTGTCTTGTGATGGAATAT-GCTGAAGGGGCTCTTTATATAATG 118  
DB 549 CTGCTTGAATCCAGTGTCTTGTGATGGAATATAGTGAAGGGGCTCTTTATATAATG 608  
QY 119 TGCTGATGGTGTGAACCATTCATATATATCTGCTGCCACGCAATGAGTTGGTGT 178



Fax: 301 838 0208  
Email: johnq@cigr.org  
Plate: 145  
Seq primer: Reverse.

## FEATURES

Location/Qualifiers  
1..542  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGF"  
/note="Vector: pBluescriptSKm"

BASE COUNT  
ORIGIN

149 a 102 c 136 g 155 t

Query Match 69.1%; Score 472.6; DB 10; Length 542;  
Best Local Similarity 99.0%; Pred. No. 9.3e-131;  
Matches 486; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGGCC 60  
Db 51 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGGCC 110  
QY 61 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 120  
Db 111 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 170  
QY 121 CTGCATGTGCTGAACCATTTGCCATATTATCTGCTGCCACGCAATGAGTTGGTGTGA 180  
Db 171 CTGCATGTGCTGAACCATTTGCCATATTATCTGCTGCCACGCAATGAGTTGGTGTGA 230  
QY 181 CAGTGTTCACAGGAGTGGCTTATCTTCACAGATGCAACCCAAAGCGCTAAATTCACAGG 240  
Db 231 CAGTGTTCACAGGAGTGGCTTATCTTCACAGATGCAACCCAAAGCGCTAAATTCACAGG 290  
QY 241 GACCTGAACACCAACAACTTACTGCTGTGTCAGGGGGGACAGTCTTAAAAATTTGTGAT 300  
Db 291 GACCTGAACACCAACAACTTACTGCTGTGTCAGGGGGGACAGTCTTAAAAATTTGTGAT 350  
QY 301 TTTGGTACAGCTGTGCAATTCAGACACACATGACCAATTAACAGGGAGTGTGCTTGG 360  
Db 351 TTTGGTACAGCTGTGCAATTCAGACACACATGACCAATTAACAGGGAGTGTGCTTGG 410  
QY 361 ATGGCACTGAAGTTTGTGAGTAGTAATTAACAGTGAATAATGACGTCCTTCAGCTGG 420  
Db 411 ATGGCACTGAAGTTTGTGAGTAGTAATTAACAGTGAATAATGACGTCCTTCAGCTGG 470  
QY 421 GGTATTATCTTTGGGAAGTGATAACCGCTC-GGAAACCCCTTTGATGAGATTGGTGGCCC 479  
Db 471 GGTATTATCTTTGGGAAGTGATAACCGCTGCGGGAACCCCTGATGAGATTGGTGGCCC 530  
QY 480 AGCTTTCCGAA 490  
Db 531 AACTTTCCCAA 541

RESULT 11  
AL550589  
LOCUS 1118 bp mRNA linear EST 31-MAY-2003  
DEFINITION AL550589 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CSODI058Y014 5-PRIME, mRNA sequence.  
ACCESSION AL550589  
VERSION AL550589.2 GI:31272406  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1118)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT On Feb 15, 2001 this sequence version replaced gi:12887700.  
Contact: Genoscope

Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3613.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSODI058BH070P1&cluster=3613.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paraday Avenue Genoscope sequence ID : CSODI058BH07Q1.

## FEATURES

Location/Qualifiers  
1..1118  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="CSODI058Y014"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 276 a 227 c 292 g 274 t 49 others  
ORIGIN

Query Match 68.5%; Score 468.2; DB 9; Length 1118;  
Best Local Similarity 94.3%; Pred. No. 3.1e-129;  
Matches 510; Conservative 7; Mismatches 20; Indels 4; Gaps 3;

QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGGCC 60  
Db 454 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGGCC 513  
QY 61 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 120  
Db 514 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 573  
QY 121 CTGCATGTGCTGAACCATTTGCCATATTATCTGCTGCCACGCAATGAGTTGGTGTGA 180  
Db 574 CTGCATGTGCTGAACCATTTGCCATATTATCTGCTGCCACGCAATGAGTTGGTGTGA 633  
QY 181 CAGTGTTCACAG-GAGTGGCTTATCTTCACAGATGCAACCCAAAGCGCTAAATTCACAG 239  
Db 634 CAGTGTTCACAGAGAGTGGCTTATCTTCACAGATGCAACCCAAAGCGCTAAATTCACAG 693  
QY 240 GGACCTGAACACCAACCACTTACTGCTGGTTGAGGGGGACAGTCTAAAAATTTGTGA 299  
Db 694 GGACCTGAACACCAACCACTTACTGCTGGTTGAGGGGGACAGTCTAAAAATTTGTGA 753  
QY 300 TTTTGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTGCTTG 359  
Db 754 TTTTGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTGCTTG 813  
QY 360 GATGGCACCTGGAAGTTTTCAGAGTAGTAATTCAGTGAATAATGTCAGCTTTCAGCTG 419  
Db 814 GATGGCACCTGGAAGTTTTCAGAGTAGTAATTCAGTGAATAATGTCAGCTTTCAGCTG 873  
QY 420 GGGTATTATTTCTTTGGGAAGTGAACCGCTCGGAAACCCCTTTGATGAGATTGGTGGCCC 479  
Db 874 GGGTATTATTTCTTTGGGAAGTGAAT-AMSCGTGGAAACCCCTTTGATGAGATTGGTGGCCC- 931  
QY 480 AGCTTTCCGAATCATGTGGCTCTTATATATGTTACTCGACCACTGATAAAAAATTT 539  
Db 932 -AGCTTTCSAATCATGTGGGCTGTTTATATATGTTATCGACMACACTGATAAAAAATTCCT 990  
QY 540 A 540  
Db 991 A 991

RESULT 12  
BU418819





ACCESSION CB583299  
VERSION CB583299.1 GI:29528785  
KEYWORDS EST.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
REFERENCE 1 (bases 1 to 590)  
AUTHORS Angen EST Program.  
TITLE Angen Rat EST Program  
JOURNAL Unpublished  
COMMENT Contact: Dan Fitzpatrick  
Angen, Inc  
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA  
Tel: 805 447-4881  
Plate: 00366 row: h column: 12.  
FEATURES  
source  
1..590  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/clone="nrhy5-00366-h12"  
/clone\_lib="W Rat hypothalamus (10471)"  
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat  
hypothalamus adult female Wistar rat avg. insert size 2.3  
kb fraction 6 and 7"  
BASE COUNT 162 a 141 c 144 g 143 t  
ORIGIN  
Query Match 63.4%; Score 433.6; DB 14; Length 590;  
Best Local Similarity 91.3%; Pred. No. 5.3e-119;  
Matches 460; Conservative 0; Mismatches 44; Indels 0; Gaps 0;  
QY 181 CAGTGTCCACAGGAGTGGCTTATCTTCACAGCATCCACCCAAAGCGTAAATTCACAGG 240  
Db |||||||  
2 CAGTGTCCACAGGAGTGGCTTACCTGCACAGCATCCACCCAAAGCGTAAATTCACAGG 61  
QY 241 GACCTGAACACCAACCTTACTGCTGGTTGACAGGGGACAGTTCCTAAATTTCTGAT 300  
Db |||||||  
62 GACCTGAACCTCCAACTTGTCTGGTTGACAGGGGACAGTTCCTAAATTTCTGAT 121  
QY 301 TTGTGTACAGCTGTGACATTCAGACACATGACCAATTAAGGGAGTCTGCTTGG 360  
Db |||||||  
122 TTGTGTACAGCTGTGACATTCAGACACATGACCAATTAAGGGAGTCTGCTTGG 181  
QY 361 ATGGCACCTGAAGTTTTCAGGTAGTAAATACAGTGAATATGACGCTTCAGCTGG 420  
Db |||||||  
182 ATGGCACCGAAGTGTTCAGGTAGTAAATACAGTGAATATGACGCTTCAGCTGG 241  
QY 421 GGTATTATCTTTGGGAAGTATGACGCTCGGAAACCTTTGATGAGATTGGTGCCCA 480  
Db |||||||  
242 GGTATTATCTTTGGGAAGTATGACGCTCGGAAACCTTTGATGAGATTGGTGCCCA 301  
QY 481 GCTTTCCGAATCATGTGGCTGTTTCAATATGATCTCGACACCTCTGATAAAAAATTGA 540  
Db |||||||  
302 GCTTTCCGAATCATGTGGCTGTTTCAATATGATCTCGACACCTCTGATAAAAAATTG 361  
QY 541 CCTAAGCCCATTCAGAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db |||||||  
362 CCTAAGCCCATTCAGAGCTGATGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421  
QY 601 TCAATGGAGGAATTTGGAATAATGATCTCATTGATCGGCTACTTTCCAGGACAGAT 660  
Db |||||||  
422 TCAATGGAGGAATTTGGAATAATGATCTCATTGATCGGCTACTTTCCAGGACAGAT 481  
QY 661 GAGCATTACAGTATCCTTGTCTCAG 684  
Db |||||||  
482 GAGCGTTACAGTATCCTTGTCTCAG 505

Search completed: December 4, 2003, 07:04:36  
Job time : 2458.23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 09:34:12 ; Search time 118.541 Seconds  
(without alignments)  
1548.515 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSTLTQLQSTNTHQSSSS.....AEFYRLMSVDHGEQSVVTP 68

Scoring table:

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Xgapop 6.0 , Xgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xhl  
-Q=/cgn2.1/USPTO.spool/US09830144/runat 03122003 122509 23313/app query fasta\_1.654  
-DB=N Geneseq 19Jun03 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144@cgn 1.0 @runat 03122003 122509 23313 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03:\*

1:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25:	/SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	359	100.0	1515	21	AAZ48861	Human TAB1 coding
2	359	100.0	1560	18	AAT91175	Human TAB1 (TAK1 b
3	359	100.0	1560	18	AAT91178	Human TAB1 (TAK1 b
4	359	100.0	1560	20	AA556278	Human TAB1 encodin
5	359	100.0	1560	21	AAA39106	Human TAB-1 nucleo
6	359	100.0	1568	20	AA556310	Human TAB1 encodin
7	359	100.0	1569	20	AA556282	Human TAB1-FLAG en
8	359	100.0	3450	25	ABX34568	Human mdt cdNA SE
9	359	100.0	16877	22	ABA20494	Human nervous syst
10	359	100.0	16877	22	AA136984	Human musculoskele
11	359	100.0	16877	25	ABX59972	cdNA encoding nove
12	352	98.1	696	21	AAF15895	Human prostate can
13	76	21.2	519	24	ABN61139	Human cancer relat
14	75	20.9	14091	23	ABL11586	Drosophila melanog
15	74.5	20.8	990	23	AA580813	DNA encoding novel
16	71.5	19.9	1108	14	AAQ49396	Class II AP endonu
17	71	19.8	996	24	ABK78928	Bacillus clausii g
18	70	19.5	3489	23	ABL28043	Drosophila melanog
19	70	19.5	19674	23	ABL28042	Drosophila melanog
20	69.5	19.4	1492	22	AAH78731	Human H4SG84 seri
21	69.5	19.4	3820	23	ABL12378	Drosophila melanog
22	69.5	19.4	4877	23	ABL13708	Drosophila melanog
23	69.5	19.4	9507	22	AA107097	Human reproductive
24	69.5	19.4	9507	25	ABZ73900	Secreted protein g
25	69.5	19.4	9507	25	ABZ67483	Human secreted pro
26	69	19.2	6743	23	AA565172	DNA encoding novel
27	69	19.2	6743	23	AA588306	DNA encoding novel
28	69	19.2	12409	22	AA303238	DNA encoding rena
29	69	19.2	23580	22	AAK66230	Human immune/haema
30	69	19.2	23580	22	AAK83578	Human immune/haema
31	68.5	19.1	3068	22	AA514723	Human phosphoinosi
32	68.5	19.1	3082	22	AA514724	Human phosphoinosi
33	68.5	19.1	12094	23	AA59536	Protonibacterium
34	68.5	19.1	16951	22	AAH48620	Human fascin DNA f
35	68.5	19.1	16951	22	AAH48622	Human fascin DNA f
36	68	18.9	273	22	ABA70748	Human foetal liver
37	68	18.9	273	22	ABA37249	Probe #15715 for g
38	68	18.9	273	22	AAK19001	Human bone marrow
39	68	18.9	273	22	AAK44951	Probe #15009 for g
40	68	18.9	273	22	AAI25076	Probe #19607 used
41	68	18.9	273	22	AAI50921	Probe #19607 used
42	68	18.9	273	23	ABS44616	Human liver single
43	68	18.9	273	24	ABS19195	Human genome-deriv
44	68	18.9	8918	22	AA526708	Human genomic DNA
45	68	18.9	8918	25	ABX74057	Human novel polynu

# ALIGNMENTS

RESULT 1

AAZ48861

ID AAZ48861 standard; cDNA; 1515 BP.

XX AAZ48861;

DT 24-MAR-2000 (first entry)

XX Human TAB1 coding sequence.

XX Human; TAB1; XIAP; X-linked inhibitor of apoptosis protein; TGF-beta;  
transforming growth factor-beta activated kinase 1; monocyte migration;  
TAK1 binding protein 1; extracellular matrix protein production;  
cell growth inhibitor; beta-amyloid protein deposition;  
immunosuppression; Transforming growth factor-beta; ds.

XX Homo sapiens.

OS

XX JPL1326328-A.  
 PN  
 XX  
 XX  
 PD 26-NOV-1999.  
 XX  
 PF 13-MAY-1998; 98JP-0130378.  
 XX  
 PR 13-MAY-1998; 98JP-0130378.  
 XX  
 XX (MATSU) MATSUMOTO K.  
 PA  
 XX WPI; 2000-078337/07.  
 XX P-PSDB; AAY59450.  
 DR  
 XX  
 XX Screening a substance which inhibits combination of the X-linked  
 PT inhibitor of apoptosis protein -  
 XX  
 XX Claim 2; Page 25-26; 43pp; Japanese.  
 PS  
 XX This sequence encodes the human TAB1 protein.  
 CC The invention relates to a method for screening a substance inhibiting  
 CC the formation of a complex between XIAP and TAB1, in which X-linked  
 CC inhibitor of apoptosis protein (XIAP), transforming growth factor-beta  
 CC activated kinase 1 (TAK1) binding protein 1 (TAB1) and a substance to be  
 CC tested are contacted with each other and then the presence or formation  
 CC of a complex between XIAP and TAB1 is detected. The substance can be used  
 CC as a drug for extracellular matrix protein production enhancement, cell  
 CC growth inhibition, monocyte migration, physiologically active substance  
 CC induction, immunosuppression, and beta-amyloid protein deposition. A  
 CC substance inhibiting the formation of a complex between TAB1 and XIAP as  
 CC well as between XIAP and TGF-beta (Transforming growth factor-beta) type  
 CC I and/or type II receptor is useful as a drug.  
 XX  
 SQ Sequence 1515 BP; 323 A; 457 C; 463 G; 272 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,32e-35 Length: 1515  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x AA248861 (1-1515)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 DB 1309 CAAAGCCGACCTTAACCCCTGCAGTCCACCAACGACGACGAGCAGCTCCAGC 1368

QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 DB 1369 TCTGACGAGGAGCTTCTCCGCTCCGCGCCGCGCCACCTCGCTCCGCGGAGAGCGT 1428

QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 DB 1429 CGTGTTCAGCCCTATGTGGACTTGTGAGTTTACCGCTCTCGACGCGGACCATGCC 1488

QY 61 GluGlnSerValThrAlaPro 68  
 DB 1489 GAGCAGAGCGTGGTGACAGCACCG 1512

RESULT 2  
 AAT91175  
 ID AAT91175 standard; cDNA; 1560 BP.  
 XX  
 AC AAT91175;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 14-APR-1998 (first entry)  
 XX  
 DE Human TAB1 (TAK1 binding protein) cDNA.  
 XX  
 XX TAB1; TAK1 binding protein; transforming growth factor-beta;  
 KW

KW signal transduction; human; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 30..1544  
 FT /\*tag= a  
 FT variation 185  
 FT /\*tag= b  
 FT /note= "another clone has adenine at position 185,  
 FT with codon AGC (Ser) altered to AGA (Arg)";  
 XX  
 PN EP803571-A2.  
 XX  
 PD 29-OCT-1997.  
 XX  
 PF 24-APR-1997; 97EP-0302808.  
 XX  
 PR 20-NOV-1996; 96US-0752891.  
 PR 24-APR-1996; 96JP-0126282.  
 PR 28-OCT-1996; 96JP-0300856.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Matsumoto K, Nishida E;  
 XX  
 XX WPI; 1997-515318/48.  
 DR P-PSDB; AAW26706.  
 XX  
 XX DNA encoding TAK1 binding protein TAB1 - member of transforming  
 PT growth factor beta receptor signal production pathway, which  
 PT activates TAK-1 kinase activity upon binding  
 XX  
 PS Claim 1; Page 17-19; 30pp; English.  
 XX

This cDNA clone codes for human TAB1 (see AAW26706), a novel member of the transforming growth factor-beta receptor signal transduction pathway, which activates TAK-1 kinase activity upon binding. To obtain the full-length TAB1 coding sequence, a human kidney cDNA library was screened using as a probe a partial TAB1 cDNA obtained from a yeast two-hybrid assay for proteins that interacted with TAK1. The 5' terminus was identified by 5'RACE. 2 Different clones were sequenced, with cytosine and adenine (see AAT91178) as the 185th nucleotide, respectively, and deposited as FERM BP-5599 and FERM BP-5508, respectively. Also claimed are: isolated DNA encoding a protein modified by a substitution, deletion and/or addition of 1 or more amino acids of the 504-residue TAB1 sequence; (2) DNA which can hybridize with the 1560 bp nucleic acid sequence; (3) isolated DNA encoding a protein comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence; (4) DNA encoding a fusion protein comprising an above protein or polypeptide; (5) expression vector comprising an above DNA; and (6) host cell, preferably a mammalian or yeast cell, transformed by the expression vector. Cells expressing TAB1 and TAK1 can be used to screen for TGF-beta signalling pathway inhibitors by contacting the cells with a test compound, and measuring the TAK1 kinase activity.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX

SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,41e-35 Length: 1560  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 18 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x AAT91175 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 DB 1338 CAAAGCCGACCTTAACCCCTGCAGTCCACCAACGACGACGAGCAGCTCCAGC 1397



QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1398 TCTGACGAGGACCTTCCGCTCCGCGCCGCCCACTCGCTCCGCGCTGGGAGACGGT 1457  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1458 CGTGTGAGCCCTATGTGGACTTGTCTGAGTTTACCGCCTCTGGAGGCTGGACCATGGC 1517  
 QY 61 GluGlnSerValThrAlaPro 68  
 Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 3  
 AAT91178  
 ID AAT91178 standard; cDNA; 1560 BP.  
 XX AC AAT91178;  
 XX DT 25-MAR-2003 (updated)  
 XX DT 14-APR-1998 (first entry)  
 XX DE Human TAB1 (TAK1 binding protein) cDNA.  
 XX KW TAB1; TAK1 binding protein; transforming growth factor-beta;  
 KW signal transduction; human; ds.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 30..1544  
 FT /\*tag= a  
 FT variation 185  
 FT /\*tag= b  
 FT /note= "another clone has cytosine at position 185,  
 with codon AGA (Arg) altered to AGC (Ser)"  
 PN EP803571-A2.  
 XX 29-OCT-1997.  
 XX 24-APR-1997; 97EP-0302808.  
 XX 20-NOV-1996; 96US-0752891.  
 XX 24-APR-1996; 96JP-0126282.  
 XX 28-OCT-1996; 96JP-0300856.  
 XX (CHUS ) CHUGAI SEIYAKU KK.  
 XX Matsumoto K, Nishida E;  
 XX WPI; 1997-515318/48.  
 XX P-PSDB; AAW26707.  
 PT DNA encoding TAK1 binding protein TAB1 - member of transforming  
 PT growth factor beta receptor signal production pathway, which  
 PT activates TAK-1 kinase activity upon binding  
 XX Example 5; Page 19-21; 30pp; English.  
 CC This cDNA clone codes for human TAB1 (see AAW26707), a novel member  
 CC of the transforming growth factor-beta receptor signal transduction  
 CC pathway, which activates TAK-1 kinase activity upon binding. To  
 CC obtain the full-length TAB1 coding sequence, a human kidney cDNA  
 CC library was screened using as a probe a partial TAB1 cDNA obtained  
 CC from a yeast two-hybrid assay for proteins that interacted with  
 CC TAK1. The 5' terminus was identified by 5'RACE. 2 Different  
 CC clones were sequenced, with cytosine (see AAT91175) and adenine as  
 CC the 185th nucleotide, respectively, and deposited as FERM BP-5599  
 CC and FERM BP-5508, respectively. Also claimed are: isolated DNA  
 CC encoding a protein modified by a substitution, deletion and/or  
 CC addition of 1 or more amino acids of the 504-residue TAB1 sequence;  
 CC (2) DNA which can hybridise with the 1560 bp nucleic acid sequence;

CC (3) isolated DNA encoding a protein comprising amino acids 21-579  
 CC or 437-504 of the 504 TAB1 sequence; (4) DNA encoding a fusion  
 CC protein comprising an above protein or polypeptide; (5) expression  
 CC vector comprising an above DNA; and (6) host cell, preferably a  
 CC mammalian or yeast cell, transformed by the expression vector.  
 CC Cells expressing TAB1 and TAK1 can be used to screen for TGF-beta  
 CC signalling pathway inhibitors by contacting the cells with a test  
 CC compound, and measuring the TAK1 kinase activity.  
 CC (Updated on 25-MAR-2003 to correct FA field.)  
 XX SQ Sequence 1560 BP; 333 A; 468 C; 480 G; 279 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 2.41e-35 Length: 1560  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 18 Gaps: 0  
 US-09-830-144-4\_COPY\_437\_504 (1-68) x AAT91178 (1-1560)  
 QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1338 CAAAGCCCGACCTTAACCTGCTCCACCAACACGACACGACGAGCAGCTCCAGC 1397  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1398 TCTGACGAGGACCTTCCGCTCCGCGCCGCCCACTCGCTCCGCGCTGGGAGACGGT 1457  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1458 CGTGTGAGCCCTATGTGGACTTGTCTGAGTTTACCGCCTCTGGAGGCTGGACCATGGC 1517  
 QY 61 GluGlnSerValThrAlaPro 68  
 Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 4  
 AAX56278  
 ID AAX56278 standard; DNA; 1560 BP.  
 XX AC AAX56278;  
 XX DT 21-JUL-1999 (first entry)  
 XX DE Human TAB1 encoding DNA.  
 XX KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
 KW transforming growth factor beta; ss.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 30..1544  
 FT /\*tag= a  
 XX WO9921010-A1.  
 XX 29-APR-1999.  
 XX 22-OCT-1998; 98WO-JP04796.  
 XX 22-OCT-1997; 97JP-0290188.  
 XX (CHUS ) CHUGAI SEIYAKU KK.  
 XX Ohtomo T, Ono K, Tsuchiya M;  
 XX WPI; 1999-312645/26.  
 XX P-PSDB; AAY09541.  
 XX Screening for TGF-beta inhibitory substances, which are useful as

PT drugs for treatment of diseases relating to its disorder  
 XX  
 PS Example 1; Page 143-147; 195pp; Japanese.  
 XX

CC A method has been developed for screening for substances which inhibit  
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method  
 CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 CC and (b) detecting the amount of bound polypeptide, in which the sample  
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming  
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
 CC indications e.g. as TGF-beta signal transduction inhibitors or  
 CC activators, or extracellular matrix protein production enhancement  
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or  
 CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence encodes human TAB1.

SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,41e-35 Length: 1560  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x AAX56278 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 DB 1338 CAAAGCCGACCTTAACCTCTGAGTCCACCAACGACGACGAGCAGCTCCAGC 1397  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 DB 1398 TCTGACGAGGCTCTTCGCTCCGCGCCGCGCCACTCGCTCCGCTGCGGAGCGGT 1457  
 QY 41 ArgValGluProTrpValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 DB 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCAATGC 1517  
 QY 61 GluGlnSerValValThrAlaPro 68  
 DB 1518 GAGCAGAGCGTGTGACAGCACCG 1541

RESULT 5  
 AAA39106  
 ID AAA39106 standard; DNA; 1560 BP.  
 XX  
 AC AAA39106;

XX 04-SEP-2000 (first entry)  
 XX Human TAB-1 nucleotide sequence SEQ ID NO:3.  
 DE

KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;  
 KW screening; signal transduction; inhibition; inflammatory cytokine;  
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;  
 KW antiinflammatory; suppression; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH 30..1544  
 CDS /\*tag= a  
 FT /\*product= "TAB-1"  
 FT

XX WC2000023610-A1.

XX 27-APR-2000.

PD

XX 21-OCT-1999; 99WO-JP05817.  
 XX  
 PR 21-OCT-1998; 98JP-0299962.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;  
 XX WPI; 2000-339707/29.  
 DR P-PSDB; AAY91001.  
 XX  
 PT Method for screening inhibitors of TAK1 signal transduction for  
 PT suppression of inflammatory cytokine production and use as  
 PT antiinflammatory agents -  
 XX  
 PS Disclosure; Page 85-90; 100pp; Japanese.

CC The present invention describes a method for screening compounds for  
 CC inhibition of inflammatory cytokine signal transduction by contacting  
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition  
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds  
 CC for inhibition of inflammatory cytokine signal transduction in which the  
 CC inhibition of TAK1 phosphorylation is selected for; and drug  
 CC compositions for the treatment of inflammatory disorders containing as  
 CC active component an inflammatory cytokine signal transduction inhibitor.  
 CC TAK1 is an essential component of the signalling process which results  
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),  
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used  
 CC for the selection of effective antiinflammatory agents. The present  
 CC sequence encodes human TAB-1, which is used in the exemplification of the  
 CC present invention.

XX SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Alignment Scores:  
 Pred. No.: 2,41e-35 Length: 1560  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x AAA39106 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 DB 1338 CAAAGCCGACCTTAACCTCTGAGTCCACCAACGACGACGAGCAGCTCCAGC 1397  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 DB 1398 TCTGACGAGGCTCTTCGCTCCGCGCCGCGCCACTCGCTCCGCTGCGGAGCGGT 1457  
 QY 41 ArgValGluProTrpValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 DB 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCAATGC 1517  
 QY 61 GluGlnSerValValThrAlaPro 68  
 DB 1518 GAGCAGAGCGTGTGACAGCACCG 1541

RESULT 6  
 AAX56310  
 ID AAX56310 standard; DNA; 1568 BP.  
 XX  
 AC AAX56310;

XX 21-JUL-1999 (first entry)

XX Human TAB1 encoding DNA SEQ ID NO:42.

XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
 KW transforming growth factor beta; ss.

```
XX OS Homo sapiens.
XX AC AAX56282;
XX FH Key Location/Qualifiers
XX FT 11..1552
XX FT CDS /*tag= a
XX XX
XX PN WO9921010-A1.
XX XX
XX PD 29-APR-1999.
XX PF 22-OCT-1998; 98WO-JP04796.
XX PF 22-OCT-1997; 97JP-0290188.
XX XX
XX XX (CHUS ) CHUGAI SEIYAKU KK.
XX PI Ohtomo T, Ono K, Tsuchiya M;
XX XX WPI; 1999-312645/26.
XX DR P-PSDB; AAY09550.
XX XX
XX FT Screening for TGF- beta inhibitory substances, which are useful as
XX FT drugs for treatment of diseases relating to its disorder
XX XX
XX PS Example 13; Page 182-186; 195pp; Japanese.
XX CC
XX CC A method has been developed for screening for substances which inhibit
XX CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX CC comprises: (a) contacting the polypeptide in the presence of a sample;
XX CC and (b) detecting the amount of bound polypeptide, in which the sample
XX CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX CC indications e.g. as TGF-beta signal transduction inhibitors or
XX CC activators, or extracellular matrix protein production enhancement
XX CC inhibitors, or monocyte migration inhibitors or activators, or
XX CC physiological activity induction inhibitors or activators, or
XX CC immunosuppression inhibitors or activators, or amyloid beta protein
XX CC precipitation inhibitors or activators, and such substances can also be
XX CC inhibitors of the TAK1 polypeptide function, particularly kinase
XX CC activity. The present sequence encodes human TAB1.
XX SQ Sequence 1568 BP; 339 A; 472 C; 477 G; 280 T; 0 other;

Alignment Scores:
Pred. No.: 2.43e-35 Length: 1568
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AAX56310 (1-1568)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1346 CAAGCCCGACCTTAACCCCTGCAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1405
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 1406 TCTGACGAGGCGCTCTTCGCTCCCGCCGCGCCACTCGCTCCCGCTGGCAGACCGT 1465
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1466 CGTGTGTGAGCCATATGTGACCTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 1525
QY 61 GluGlnSerValValThrAlaPro 68
Db 1526 GAGCAGAGCGTGTGACAGCACCG 1549

RESULT 7
AAX56282
```

```
ID XX AAX56282 standard; DNA; 1569 BP.
XX AC AAX56282;
XX XX
XX DT 21-JUL-1999 (first entry)
XX XX
XX DE Human TAB1-FLAG encoding DNA.
XX XX
XX KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;
XX KW transforming growth factor beta; ss.
XX XX
XX OS Homo sapiens.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 7..1560
XX FT /*tag= a
XX XX
XX PN WO9921010-A1.
XX XX
XX PD 29-APR-1999.
XX PF 22-OCT-1998; 98WO-JP04796.
XX PF 22-OCT-1997; 97JP-0290188.
XX XX
XX XX (CHUS ) CHUGAI SEIYAKU KK.
XX XX
XX XX Ohtomo T, Ono K, Tsuchiya M;
XX XX WPI; 1999-312645/26.
XX XX P-PSDB; AAY09546.
XX XX
XX PT Screening for TGF- beta inhibitory substances, which are useful as
XX PT drugs for treatment of diseases relating to its disorder
XX XX
XX PS Example 1; Page 159-163; 195pp; Japanese.
XX CC
XX CC A method has been developed for screening for substances which inhibit
XX CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX CC comprises: (a) contacting the polypeptide in the presence of a sample;
XX CC and (b) detecting the amount of bound polypeptide, in which the sample
XX CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX CC indications e.g. as TGF-beta signal transduction inhibitors or
XX CC activators, or extracellular matrix protein production enhancement
XX CC inhibitors, or monocyte migration inhibitors or activators, or
XX CC physiological activity induction inhibitors or activators, or
XX CC immunosuppression inhibitors or activators, or amyloid beta protein
XX CC precipitation inhibitors or activators, and such substances can also be
XX CC inhibitors of the TAK1 polypeptide function, particularly kinase
XX CC activity. The present sequence encodes TAB1-FLAG from an example of
XX CC the present invention.
XX SQ Sequence 1569 BP; 343 A; 466 C; 476 G; 284 T; 0 other;

Alignment Scores:
Pred. No.: 2.43e-35 Length: 1569
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AAX56282 (1-1569)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1315 CAAGCCCGACCTTAACCCCTGCAGTCCACCAACAGCAGCAGCAGCTCCAGC 1374
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
```

Db 1375 TCTGACGAGGCGCTCTTCGCTCCGCGCGCCGCACTCGCTCCGCGCTGGGAGGACGGT 1434

Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
|||||

Db 1435 CGTGTTCAGCGCCATATGTGGACTTTGTCTGAGTTTACCGCTCTGTGGAGCGTGGACCATGGC 1494

Qy 61 GluGlnSerValValThrAlaPro 68  
|||||

Db 1495 GAGCAGAGCGTGGTGACAGCACCG 1518

RESULT 8  
ABX34568

ID ABX34568 standard; cDNA; 3450 BP.

XX AC ABX34568;

XX 13-FEB-2003 (first entry)

XX DE Human mddt cDNA SEQ ID 129.

XX KW MDDT; human; disease detection and treatment molecule polypeptide;  
KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;  
KW haemostatic; nephrotropic; antianaemic; antiparasitic; hepatotropic;  
KW gene therapy; protein replacement therapy; cell proliferative disorder;  
KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;  
KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;  
KW psoriasis; hepatitis; gene; ss.

XX OS Homo sapiens.

XX KW

XX W0200279449-A2.

XX 10-OCT-2002.

XX 27-MAR-2002; 2002WO-US09944.

XX 28-MAR-2001; 2001US-279619P.

XX 29-MAR-2001; 2001US-280067P.

XX 29-MAR-2001; 2001US-280068P.

XX 16-MAY-2001; 2001US-291280P.

XX 17-MAY-2001; 2001US-291829P.

XX 17-MAY-2001; 2001US-291849P.

XX 19-JUN-2001; 2001US-299428P.

XX 20-JUN-2001; 2001US-299776P.

XX 20-JUN-2001; 2001US-300001P.

XX (INCY-) INCYTE GENOMICS INC.

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;  
PI Daugherty SC, Dam TC, Liu IF, Nguyen DA, Kleefeld Y, Gerstin EH;  
PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
PI Flores V, Marwaha R, Lo A, Lan RY, Urashka MF;  
XX WPI: 2003-058431/05.  
XX P-PSDB; ABU11578.

XX New purified disease detection and treatment molecule proteins and  
PT polynucleotides, useful for diagnosing, treating or preventing cancers  
PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis  
PT or hepatitis

XX Claim 1; SEQ ID NO 129; 339pp + Sequence Listing; English.

XX This invention describes a novel disease detection and treatment molecule  
CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,  
CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,  
CC antianaemic, antiparasitic and hepatotropic activity. The polynucleotides  
CC and the polypeptides of the invention can be used for gene therapy,  
CC protein replacement therapy and are useful for treating a variety of  
CC diseases or conditions. These polypeptides or polynucleotides are  
CC particularly useful for diagnosing, treating or preventing cell

CC proliferative disorders (e.g. cancers including adenocarcinoma,  
CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's  
CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's  
CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or  
CC hepatitis. ABX34440-ABX34835 encode the MDDT polypeptides represented in  
CC ABU11450-ABU11845, described in the disclosure of the invention.  
CC NOTE: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 3450 BP; 711 A; 1039 C; 1042 G; 658 T; 0 other;

Alignment Scores:

Pred. No.:	6.86e-35	Length:	3450
Score:	359.00	Matches:	68
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	25	Gaps:	0

US-09-830-144-4\_COPY\_437\_504 (1-68) x ABX34568 (1-3450)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
|||||

Db 1345 CAAAGCCGACCTTACCTTGAGTCCACCAACAGCACGACGACGACGCTCCAGC 1404  
|||||

Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
|||||

Db 1405 TCTGACGAGGCGCTCTTCGCTCCGCGCGCCGCACTCGCTCCGCGCTGGGAGGACGGT 1464  
|||||

Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
|||||

Db 1465 CGTGTTCAGCGCCATATGTGGACTTTGTCTGAGTTTACCGCTCTGTGGAGCGTGGACCATGGC 1524  
|||||

Qy 61 GluGlnSerValValThrAlaPro 68  
|||||

Db 1525 GAGCAGAGCGTGGTGACAGCACCG 1548  
|||||

RESULT 9  
ABX20494

ID ABX20494 standard; DNA; 16877 BP.

XX AC ABX20494;

XX 23-JAN-2002 (first entry)

XX DE Human nervous system related polynucleotide SEQ ID NO 12825.

XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antiskilling; antianaemic; antiarthritic; cancer;  
KW antihemetic; hepatotropic; cerebroprotective; antinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX OS Homo sapiens.

XX PN W0200159063-A2.

XX 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US01334.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0215647.  
PR 07-JUL-2000; 2000US-0215680.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

Nucleic acids encoding 324 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases -

Disclosure; SEQ ID NO 12825; 1701pp + Sequence Listing; English.

The invention relates to novel genes (ABA11004-ABA21534) and proteins (ABB14678-ABB18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast

CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune;  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published pct sequences](http://ftp.wipo.int/pub/published/pct/sequences).

Sequence 16877 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other;  
SQ

Alignment Scores:		
Pred. No.:	5.57e-34	16877
Score:	359.00	68
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	22	Gaps: 0

US-09-830-144-4 COPY 437 504 (1-68) x ABA20494 (1-16877)

Qy	1	GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 20
Db	15011	CAAAAGCCGACCTTAACCCCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 15070
Qy	21	SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db	15071	TCTGAGCGGAGGCTCTTCGCGTCCCGCGCCGCCCACTCGTCCCGCTGGCGAGACGGT 15130
Qy	41	ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db	15131	CGTGTTCAGCCCTATGTGGACCTTCTGTAGTTTACCGCCCTCGGAGCGCTGACCATGGC 15190
Qy	61	GluGlnSerValValThrAlaPro 68
Db	15191	GAGCAGAGCCGTGGTGCAGCAGCG 15214

## RESULT, T 10

RESULI TO  
AAT.36984

AAL36984  
ID AAL36984 standard: DNA: 16877 BP.

AA AAL36984:  
ACXX  
DT 08--JAN-2002 (first entry)

XX DE Human musculoskeletal system related polynucleotide SEO ID NO 3349.

Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 anti-allergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein;  
 KW musculoskeletal system; ds.

vv Homo sapiens.  
 OS

AA WO200155367-A1.

XX  
PD 02-AUG-2001.

17-TAN-2001: 2001WQ-US01338.

31-TAN-2000: 2000US-0179065

04-FEB-2000; 2000US=0180628.  
PR  
24-FEB-2000; 2000US=0184664.  
PR

02-MAR-2000; 2000US-0186350.  
PR  
16-MAR-2000; 2000US-0189874  
PR

PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123

PR	19-MAY-2000;	2000US-0205515.
PR	28-JUN-2000;	2000US-0209467.
PR	28-JUN-2000;	2000US-0214886.
PR	30-JUN-2000;	2000US-0215135.
PR	30-JUL-2000;	2000US-0216647.
PR	07-JUL-2000;	2000US-0216880.
PR	11-JUL-2000;	2000US-0217487.
PR	11-JUL-2000;	2000US-0217496.
PR	14-JUL-2000;	2000US-0218290.
PR	26-JUL-2000;	2000US-0220963.
PR	26-AUG-2000;	2000US-0220964.
PR	26-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225457.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226691.
PR	22-AUG-2000;	2000US-0226868.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	06-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	21-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0233238.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	13-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.

20-OCT-2000; 2000US-0241221.  
 20-OCT-2000; 2000US-0241785.  
 20-OCT-2000; 2000US-0241786.  
 20-OCT-2000; 2000US-0241787.  
 20-OCT-2000; 2000US-0241808.  
 20-OCT-2000; 2000US-0241809.  
 20-OCT-2000; 2000US-0241826.  
 01-NOV-2000; 2000US-0244617.  
 08-NOV-2000; 2000US-0246474.  
 08-NOV-2000; 2000US-0246475.  
 08-NOV-2000; 2000US-0246476.  
 08-NOV-2000; 2000US-0246477.  
 08-NOV-2000; 2000US-0246478.  
 08-NOV-2000; 2000US-0246523.  
 08-NOV-2000; 2000US-0246524.  
 08-NOV-2000; 2000US-0246525.  
 08-NOV-2000; 2000US-0246526.  
 08-NOV-2000; 2000US-0246527.  
 08-NOV-2000; 2000US-0246528.  
 08-NOV-2000; 2000US-0246532.  
 08-NOV-2000; 2000US-0246609.  
 08-NOV-2000; 2000US-0246610.  
 08-NOV-2000; 2000US-0246611.  
 08-NOV-2000; 2000US-0246613.  
 17-NOV-2000; 2000US-0249207.  
 17-NOV-2000; 2000US-0249208.  
 17-NOV-2000; 2000US-0249209.  
 17-NOV-2000; 2000US-0249210.  
 17-NOV-2000; 2000US-0249211.  
 17-NOV-2000; 2000US-0249212.  
 17-NOV-2000; 2000US-0249213.  
 17-NOV-2000; 2000US-0249214.  
 17-NOV-2000; 2000US-0249215.  
 17-NOV-2000; 2000US-0249216.  
 17-NOV-2000; 2000US-0249217.  
 17-NOV-2000; 2000US-0249218.  
 17-NOV-2000; 2000US-0249244.  
 17-NOV-2000; 2000US-0249245.  
 17-NOV-2000; 2000US-0249246.  
 17-NOV-2000; 2000US-0249265.  
 17-NOV-2000; 2000US-0249297.  
 17-NOV-2000; 2000US-0249299.  
 17-NOV-2000; 2000US-0249300.  
 01-DEC-2000; 2000US-0250160.  
 01-DEC-2000; 2000US-0250391.  
 05-DEC-2000; 2000US-0251030.  
 05-DEC-2000; 2000US-0251988.  
 06-DEC-2000; 2000US-0256719.  
 08-DEC-2000; 2000US-0251479.  
 08-DEC-2000; 2000US-0251856.  
 08-DEC-2000; 2000US-0251868.  
 08-DEC-2000; 2000US-0251869.  
 08-DEC-2000; 2000US-0251989.  
 08-DEC-2000; 2000US-0251990.  
 11-DEC-2000; 2000US-0254097.  
 05-JAN-2001; 2001US-02559678.  
 (HUMA-) HUMAN GENOME SCI INC.  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-451937/48.  
 DR  
 PT Isolated polypeptide for treating, preventing and/ or prognosing  
 PT disorders related to the musculoskeletal system including  
 PT musculoskeletal cancers and also for testing and detection e.g.  
 PT diagnosis -  
 XX  
 PS Example 2; SEQ ID NO 3349; 781pp + Sequence Listing; English.  
 CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful  
 CC for preventing, treating or ameliorating medical conditions e.g. by

protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.  
 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 16877 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 5.57e-34 Length: 16877  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-830-144-4\_COPY\_437\_504 (1-68) x AAL36984 (1-16877)  
 QY 1 GlnSerProThrLeuThrLeuGlnSerThrThrAsnThrHisThrGlnSerSerSerSer 20  
 DB 15011 CAAGCCCGACCTTAACCTGCAGTCCACCAACACGACACGACGAGCAGCTCCAGC 15070  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 DB 15071 TCTGACGGAGGCGCTCTTCGGCTCCGCGCCGCCACTCGCTCCGCTGGCAGGACGGT 15130  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 DB 15131 CGTGTGAGCCCTATGTGGACTTTCTGAGTTTACCGCTCTGAGCGCTGGACCATGTC 15190  
 QY 61 GluGlnSerValValThrAlapro 68  
 DB 15191 GAGCAGCGGIGTGACAGACCG 15214  
 RESULT 11  
 ABX59972  
 ID ABX59972 standard; cDNA; 16877 BP.  
 XX  
 AC ABX59972;  
 XX  
 DT 26-FEB-2003 (first entry)  
 XX  
 DE cDNA encoding novel human musculoskeletal system antigen #2316.  
 XX  
 KW Gene; ss; musculoskeletal system antigen; cancer; metastasis;  
 KW re-vascularisation; thrombosis; arteriosclerosis; mineral content;  
 KW cardiovascular condition; wound; injury; burn; angiogenesis; ulcer;  
 KW post-operative tissue repair; limb regeneration; neuronal growth;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW AIDS-related complex; chondrocyte growth; bone regeneration;  
 KW periodontal regeneration; tissue transport; bone graft; skin aging;  
 KW keratinocyte growth; hair loss; melanocyte growth; cell proliferation;  
 KW cell growth; organ transplant; cell differentiation; body height;  
 KW weight; hair colour; eye colour; skin; percentage of adipose tissue;  
 KW pigmentation; cosmetic surgery; metabolism; biorythm; cardiac rhythm;  
 KW depression; tendency for violence; pain; reproductive capability;  
 KW hormone level; endocrine level; appetite; libido; memory; stress;  
 KW storage capability; fat content; lipid content; protein content;  
 KW carbohydrate content; vitamin content; cofactor content;  
 XX nutritional component.  
 OS Homo sapiens.

```
XX
FN US2002147140-A1.
PD 10-OCT-2002.
XX
XX 17-JAN-2001; 2001US-0764877.
XX
PR 31-JAN-2000; 2000US-179065P.
PR 04-FEB-2000; 2000US-180628P.
PR 28-JUN-2000; 2000US-214886P.
PR 07-JUL-2000; 2000US-216647P.
PR 07-JUL-2000; 2000US-216880P.
PR 11-JUL-2000; 2000US-217487P.
PR 11-JUL-2000; 2000US-217496P.
PR 14-JUL-2000; 2000US-218290P.
PR 26-JUL-2000; 2000US-220963P.
PR 26-JUL-2000; 2000US-220964P.
PR 14-AUG-2000; 2000US-224518P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-224519P.
PR 14-AUG-2000; 2000US-225267P.
PR 14-AUG-2000; 2000US-225268P.
PR 14-AUG-2000; 2000US-225270P.
PR 14-AUG-2000; 2000US-225447P.
PR 14-AUG-2000; 2000US-225757P.
PR 22-AUG-2000; 2000US-225758P.
PR 22-AUG-2000; 2000US-226868P.
PR 30-AUG-2000; 2000US-228924P.
PR 01-SEP-2000; 2000US-228928P.
PR 01-SEP-2000; 2000US-229343P.
PR 01-SEP-2000; 2000US-229344P.
PR 01-SEP-2000; 2000US-229345P.
PR 05-SEP-2000; 2000US-229509P.
PR 05-SEP-2000; 2000US-229513P.
PR 08-SEP-2000; 2000US-231413P.
PR 21-SEP-2000; 2000US-234223P.
PR 21-SEP-2000; 2000US-234274P.
PR 25-SEP-2000; 2000US-234997P.
PR 27-SEP-2000; 2000US-235834P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236327P.
PR 29-SEP-2000; 2000US-236368P.
PR 29-SEP-2000; 2000US-236369P.
PR 29-SEP-2000; 2000US-236370P.
PR 02-OCT-2000; 2000US-236802P.
PR 02-OCT-2000; 2000US-237037P.
PR 02-OCT-2000; 2000US-237038P.
PR 02-OCT-2000; 2000US-237039P.
PR 02-OCT-2000; 2000US-237040P.
PR 13-OCT-2000; 2000US-239335P.
PR 20-OCT-2000; 2000US-240960P.
PR 20-OCT-2000; 2000US-241785P.
PR 20-OCT-2000; 2000US-241809P.
PR 01-NOV-2000; 2000US-244617P.
PR 17-NOV-2000; 2000US-249299P.
PR 08-DEC-2000; 2000US-251856P.
PR 08-DEC-2000; 2000US-251868P.
PR 08-DEC-2000; 2000US-251869P.
XX
XX (ROSE/) ROSEN C A.
XX (RUBE/) RUBEN S M.
XX (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2003-128199/12.
XX
XX Isolated nucleic acid molecules encoding musculoskeletal system
XX associated polypeptides, useful for detecting disorders, e.g. cancer -
XX
XX Disclosure; SEQ ID NO 3349; 321pp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
XX sequence encoding musculoskeletal system associated polypeptides useful
XX
```

for detecting disorders, e.g., cancer or cancer metastases, in animals or humans. The nucleic acid: stimulates re-vascularisation of ischaemic tissues associated with conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions; treats wounds due to injuries, burns, post-operative tissue repair, and ulcers; stimulates angiogenesis and limb regeneration; stimulates neuronal growth; can treat and prevent neuronal damage occurring in certain disorders or neurodegenerative conditions, such as, Alzheimer's disease, Parkinson's disease, and AIDS-related complex; stimulates chondrocyte growth, thus they can be used to enhance bone and periodontal regeneration and aid in tissue transports or bone grafts; prevents skin aging due to sunburn by stimulating keratinocyte growth; prevents hair loss, since FGF family members activate hair-forming cells and promotes melanocyte growth; stimulates growth and differentiation of hematopoietic cells and bone marrow cells when used in combination with other cytokines; maintains organs before transplantation or for supporting cell culture of primary tissues; induces tissue of mesodermal origin to differentiate in early embryos; increases or decreases the differentiation or proliferation of embryonic stem cells, besides, haematopoietic lineage; modulates mammalian characteristics, such as, body height, weight, hair colour, eye colour, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery); modulates mammalian metabolism; changes (e.g., metabolic state or physical state by influencing biorhythms, cardiac rhythms, depression, tendency for violence, tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, or stress; increases or decreases storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components. This sequence encodes a novel human musculoskeletal system antigen.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at

ftp.segdata.uspto.gov/sequence.html?DocID=20020147140.

Sequence 16877 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other;

Alignment Scores:

Pred. No.:	5.57e-34	Length:	16877
Score:	359.00	Matches:	68
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	25	Gaps:	0

US-09-830-144-4\_COPY\_437\_504 (1-68) x ABX59972 (1-16877)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20

Db 15011 CAAAGCCGACCTTACCTGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 15070

Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluaspGly 40

Db 15071 TCTGACGAGGCGCTCTTCGCTCCGCGCGCGCCACCTCGCTCCGCGCGAGGACGT 15130

Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60

Db 15131 CGTGTGAGCCCTAIGTGAGCTTGTGAGTTTACGCCCTCTGGAGCGTGGACCATGGC 15190

Qy 61 GluGlnSerValValThrAlaPro 68

Db 15191 GAGCAGAGCGTGTGTGACAGCACCG 15214

RESULT 12

AAF15895

ID AAF15895 standard; cDNA; 696 BP.

XX

AC AAF15895;

XX

DT 13-MAR-2001 (first entry)

XX

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:330.

XX

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

XX



KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
KW wound; infectious disease; ss.

XX Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

XX P-PsDB; AAB56692.

PT Prostate cancer associated gene sequences, referred to as prostate  
PT cancer antigens, useful for treatment, prevention, and diagnosis of  
PT disorders such as prostate cancer -

PS Claim 1; Page 837; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated  
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
CC The prostate cancer antigens can have neuroprotective, cytostatic,  
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
CC and can be used in gene therapy. The prostate cancer antigen  
CC polynucleotides may be used for detection of prostate cancer, chromosome  
CC identification, as chromosome markers, and for numerous other diagnostic  
CC or research purposes. The prostate cancer antigens may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
CC AAB57303 represent sequences used in the exemplification of the present  
CC invention.

SQ Sequence 696 BP; 143 A; 237 C; 187 G; 120 T; 9 other;

Alignment Scores:  
Pred. No.: 6.34e-35 Length: 696  
Score: 352.00 Matches: 67  
Percent Similarity: 98.53% Conservative: 0  
Best Local Similarity: 98.53% Mismatches: 1  
Query Match: 98.05% Indels: 0  
DB: 21 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x AAF15895 (1-696)

QY 1 GlnSerProThrLeuThrGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
DB 49 CAAAGCCGACCTTAACCCCTGCAGTCCACCAACGACGACGACGAGCTCCAGC 108  
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
DB 109 TCTRACGAGGCGCTCTCCGTCGCCGGCCGCCACCTCGCTCCCGCTGGCGAGACGCT 168  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
DB 169 CGTGTGAGCCCTATGTGACATTCTGAGTTTACCCTCTGGAGCGTGACCATGGC 228  
QY 61 GluGlnSerValValThrAlaPro 68  
DB 229 GAGCAGAGCGTGGTGACAGCACCG 252

RESULT 13

ABN61139/c

ID ABN61139 standard; cDNA; 519 BP.

XX AC ABN61139;

XX DT 28-JUN-2002 (first entry)

XX Human cancer related polynucleotide SEQ ID NO 1106.  
DE Human; cytostatic; gene expression; gene mapping; tissue profiling;  
KW gene therapy; cancer; tumour; gene; ss.

XX Homo sapiens.

XX WO200214500-A2.

XX PD 21-FEB-2002.

XX 16-AUG-2001; 2001WO-US25840.

XX 16-AUG-2000; 2000US-226326P.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;  
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;  
XX WPI; 2002-241905/29.  
XX New nucleic acid for producing a polypeptide, detecting differentially  
PT expressed genes correlated with a cancerous state of a mammalian cell,  
PT and inhibiting tumor growth -  
XX Claim 1; SEQ ID NO 1106; 883pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (ABN2753-ABN33262)  
CC with cytostatic activity. The polynucleotide is used to produce a  
CC polypeptide, to detect differentially expressed genes correlated with a  
CC cancerous state of a mammalian cell and to inhibit tumour growth. The  
CC polynucleotide is used as a probe in mapping and tissue profiling. The  
CC encoded polypeptide and antibodies to the polypeptide can also be used  
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for  
CC gene therapy.

CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 519 BP; 152 A; 118 C; 114 G; 135 T; 0 other;

Alignment Scores:  
Pred. No.: 2.59 Length: 519  
Score: 76.00 Matches: 22  
Percent Similarity: 49.23% Conservative: 10  
Best Local Similarity: 33.85% Mismatches: 25  
Query Match: 21.17% Indels: 8  
DB: 24 Gaps: 3

US-09-830-144-4\_COPY\_437\_504 (1-68) x ABN61139 (1-519)

QY 5 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGly 24  
DB 368 CTCCTCCATCAGAAAGTCCAGAAACACTGCTACTGAGCAGAGTCTCATGCTG- - - - -GGCTGC 315  
QY 25 LeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGlyArgValGluPro 44  
DB 314 AGATTCTCAAGGCTCTCTCTGCACAGCTCTGCGCTGGAGTAGTGGGAGGCGCATTCCTCA 255  
QY 45 --TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGlnSer 63  
DB 254 AACAGTCTCTCTTCCCC- - - - -TGGAGTGGGGAAGGAGGTACATGAGT 210

QY	64 ValValThrAlaPro 68 :::     209 CTCATCTCAATTCTT 195
Dd	
Db	
QY	RESULT 14 ABL11586 ID ABL11586 standard; cDNA; 14091 BP.
ID	
XX	ABL11586;
AC	
XX	
DT	26-MAR-2002 (first entry)
DE	
XX	Drosophila melanogaster expressed polynucleotide SEQ ID NO 29240.
KW	Drosophila; developmental biology; cell signalling; insecticide;
KW	pharmaceutical; gene; ss.
OS	Drosophila melanogaster.
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
PF	23-MAR-2001; 2001WO-US09231.
XX	
PR	23-MAR-2000; 2000US-191637P.
XX	
PP	11-JUL-2000; 2000US-0614150.
XX	
PA	(PEKE ) PE CORP NY.
PI	Venter JC, Adams M, Li PWD, Myers EW;
Pt	WPI; 2001-656860/75.
DR	P-PSDB; ABB67483.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
XX	
PS	Claim 1; SEQ ID NO 29240; 21pp + Sequence Listing; English.
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB5737-ABB72072).
CC	The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 14091 BP; 3932 A; 3192 C; 3264 G; 3703 T; 0 other;
Alignment Scores:	
Pred. No.:	270 Length: 14091
Score:	75.00 Matches: 25
Percent Similarity:	44.44% Conservative: 7
Best Local Similarity:	34.72% Mismatches: 15
Query Match:	20.89% Indels: 2
DB:	23 Gaps: 2
US-09-830-144-4_COPY_437_504 (1-68) x ABL11586 (1-14091)	
QY	2 SerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 21       ::: 9339 TCCCAGAACCGAGCTCACCCTGGCACAAAACACACGCCGTGCATCAGCGCGGACGAGC 8998
Dd	
Db	
QY	22 AspGlyCysLeuPheArgSerArgProAlaHisSer-LeuProProGly----- 37       8999 CGGGGATCATCCGTTT-----CGCGCGGACACTCATCTTCGCGCAGTGCGCCCTGCAGC 9052
Dd	
Db	



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: December 4, 2003, 09:34:57 : Search time 1463.84 Seconds  
(without alignments)  
1900.384 Million cell updates/sec  
Title: US-09-830-144-4\_COPY\_437\_504  
Perfect score: 359  
Sequence: 1 QSPTLTQSTWTHQSSSSS.....AEFYRLMSVDHGQSVVTP 68

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_epool/US09830144/runat\_03122003\_122509\_23321/app\_query.fasta\_1.654  
-DB=GenEmbl -QPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144.qcgn 1\_1\_3608 @runat\_03122003\_122509\_23321 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_lm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_hcg\_hum.\*  
31: em\_hcg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	359	100.0	1515	6	E31041	E31041 Method for
2	359	100.0	1560	6	AR058299	AR058299 Sequence
3	359	100.0	1560	6	AR058302	AR058302 Sequence
4	359	100.0	1560	6	AR088273	AR088273 Sequence
5	359	100.0	1560	6	AR088276	AR088276 Sequence
6	359	100.0	1560	6	AR116881	AR116881 Sequence
7	359	100.0	1560	6	AR116884	AR116884 Sequence
8	359	100.0	1560	6	AR231191	AR231191 Sequence
9	359	100.0	1560	6	AR307972	AR307972 Sequence
10	359	100.0	1560	6	E14752	E14752 Human mRNA
11	359	100.0	1568	6	AR231223	AR231223 Sequence
12	359	100.0	1568	6	AR308004	AR308004 Sequence
13	359	100.0	1569	6	AR231195	AR231195 Sequence
14	359	100.0	1569	6	AR307976	AR307976 Sequence
15	359	100.0	3096	9	HSU499215	U49928 Homo sapien
16	359	100.0	3253	9	BC050554	BC050554 Homo sapi
17	359	100.0	3290	9	BC038582	BC038582 Homo sapi
18	359	100.0	69660	9	HS407F17	Z83845 Human DNA s
19	359	100.0	171718	2	AC141444	AC141444 Homo sapi
20	356	99.2	1784	10	AB088136	AB088136 Mus muscu
21	356	99.2	2943	10	BC041110	BC041110 Mus muscu
22	356	99.2	2944	10	BC027054	BC027054 Mus muscu
23	356	99.2	182516	2	AC140267	AC140267 Mus muscu
24	356	99.2	268805	2	AC127784	AC127784 Rattus no
25	356	99.2	298235	2	AC127924	AC127924 Rattus no
26	260.5	72.6	1926	5	XLU92031	U92031 Xenopus lae
27	79	22.0	251693	2	AL845323	AL845323 Mus muscu
28	78	21.7	264960	2	AC115155	AC115155 Rattus no
29	78	21.7	300700	2	AC131399	AC131399 Rattus no
30	76	21.2	519	6	AX386178	AX386178 Sequence
31	76	21.2	113028	9	AL365220	AL365220 Human DNA
32	75.5	21.0	82166	9	AC006237	AC006237 Homo sapi
33	75.5	21.0	179966	2	AC023247	AC023247 Homo sapi
34	75.5	21.0	205040	9	AC015849	AC015849 Homo sapi
35	75.5	21.0	229922	2	AC102141	AC102141 Mus muscu
36	75.5	21.0	241541	2	AC084085	AC084085 Homo sapi
37	75	20.9	14090	2	AC014059	AC014059 Drosophil
38	75	20.9	158402	3	AC008343	AC008343 Drosophil
39	75	20.9	175007	3	AC008342	AC008342 Drosophil
40	75	20.9	255652	3	AE003811	AE003811 Drosophil
41	74	20.6	149984	2	AC123505	AC123505 Rattus no
42	74	20.6	204263	2	AC117064	AC117064 Rattus no
43	74	20.6	223761	2	AC127239	AC127239 Mus muscu
44	74	20.6	228896	2	AC099452	AC099452 Rattus no
45	74	20.6	229247	10	AC098888	AC098888 Mus muscu

# ALIGNMENTS

RESULT 1

```

E31041
LOCUS       E31041               1515 bp    DNA    linear    PAT 18-JUN-2001
DEFINITION   Method for screening substance inhibiting binding to XIAP.
ACCESSION   E31041
VERSION      E31041.1  GI:13017306
KEYWORDS     JP 199326328-A/1.
SOURCE       unidentified
ORGANISM     unidentified
REFERENCE    1 (bases 1 to 1515)
AUTHORS      Kunihiro, M.
TITLE        Method for screening substance inhibiting binding to XIAP
JOURNAL      Patent: JP 199326328-A 1 26-NOV-1999;
COMMENT      KUNIHIRO MATSUMOTO
OS           Unidentified
PN           JP 199326328-A/1
PD           26-NOV-1999
PP           13-MAY-1998  JP 1998130378
PR           KUNIHIRO MATSUMOTO
PI           GOIN33/536, A61K38/00, A61K38/00, A61K38/00, A61K38/22,
PC           A61K39/395,
PC           A61K45/00, A61K45/00, A61K45/00, C07K7/06, C07K14/47, PC
PC           GOIN33/536,
PC           GOIN33/536, GOIN33/536//C12N15/09, C12P21/08, A61K37/02, A61K37/02, PC
PC           A61K37/02, A61K37/24, C12N15/00
CC           Strandedness: Double;
CC           Topology: Linear;
FH           Key      Location/Qualifiers
FT           CDS      1..1512.
                     Location/Qualifiers
                     1..1515
                     /organism="unidentified"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32644"
BASE COUNT   323 a 457 c 463 g 272 t
ORIGIN
Alignment Scores:
Pred. No.:      3,3e-30      Length:      1515
Score:          359.00      Matches:      68
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    100.00%  Indels:      0
DB:             Gaps:      6

US-09-830-144-4_COPY_437_504 (1-68) x E31041 (1-1515)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1309 CAAAGCCGACCTTAACCTCGAGTCCACCAACACGACGACGAGCAGCTCCAGC 1368
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 1369 TCTGACGAGCGCTTTCGCTCCCGCCCGCCACCTCGCTCCCGCTGGGAGGACGGT 1428
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1429 CGTGTGAGCCCTATGTGACCTTTGCTGAGTTTACCGCTCTGGAGCGTGCACCATGGC 1488
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1489 GAGCAGAGCGTGGTGACAGCACC 1512

RESULT 2
LOCUS       AR058299             1560 bp    DNA    linear    PAT 29-SEP-1999
DEFINITION   Sequence 1 from patent US 5837819.
ACCESSION   AR058299
VERSION      AR058299.1  GI:5983876
KEYWORDS

US-09-830-144-4_COPY_437_504 (1-68) x AR058299 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAAGCCGACCTTAACCTCGAGTCCACCAACACGACGACGAGCAGCTCCAGC 1397
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 1398 TCTGACGAGCGCTTTCGCTCCCGCCCGCCACCTCGCTCCCGCTGGGAGGACGGT 1457
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTATGTGACCTTTGCTGAGTTTACCGCTCTGGAGCGTGCACCATGGC 1517
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCACC 1541

RESULT 3
LOCUS       AR058302             1560 bp    DNA    linear    PAT 29-SEP-1999
DEFINITION   Sequence 5 from patent US 5837819.
ACCESSION   AR058302
VERSION      AR058302.1  GI:5983879
KEYWORDS

US-09-830-144-4_COPY_437_504 (1-68) x AR058302 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAAGCCGACCTTAACCTCGAGTCCACCAACACGACGACGAGCAGCTCCAGC 1397
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 1398 TCTGACGAGCGCTTTCGCTCCCGCCCGCCACCTCGCTCCCGCTGGGAGGACGGT 1457
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTATGTGACCTTTGCTGAGTTTACCGCTCTGGAGCGTGCACCATGGC 1517
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCACC 1541

```

```
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 1398 TCTGACGAGGCGCTCTTCGGCTCCGGCCGCCCACTCGCTCCGGCTGGCGAGACGGT 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517
QY 61 GluGlnSerValValThraPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 4
AR088273 LOCUS 1560 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from patent US 5989862.
ACCESSION AR088273
VERSION AR088273.1 GI:10015036
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1560)
AUTHORS Matsuomoto,K. and Nishida,E.
TITLE Tabi protein and DNA coding therefor
JOURNAL Patent: US 5989862-A 1 23-NOV-1999;
FEATURES
Location/Qualifiers
source 1..1560
BASE COUNT 332 a 469 c 480 g 279 t
ORIGIN
Alignment Scores:
Pred. No.: 3.4e-30 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AR088273 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAGCCCGACCTTAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCAGCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 1398 TCTGACGAGGCGCTCTTCGGCTCCGGCCGCCCACTCGCTCCGGCTGGCGAGACGGT 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517
QY 61 GluGlnSerValValThraPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 5
AR088276 LOCUS 1560 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 5 from patent US 5989862.
ACCESSION AR088276
VERSION AR088276.1 GI:10015039
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1560)
AUTHORS Matsuomoto,K. and Nishida,E.
TITLE Tabi protein and DNA coding therefor
JOURNAL Patent: US 5989862-A 5 23-NOV-1999;
FEATURES
Location/Qualifiers
```

```
source 1..1560
/organism="unknown"
BASE COUNT 333 a 468 c 480 g 279 t
ORIGIN
Alignment Scores:
Pred. No.: 3.4e-30 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AR088276 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAGCCCGACCTTAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 1398 TCTGACGAGGCGCTCTTCGGCTCCGGCCGCCCACTCGCTCCGGCTGGCGAGACGGT 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517
QY 61 GluGlnSerValValThraPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 6
AR116881 LOCUS 1560 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6140042.
ACCESSION AR116881
VERSION AR116881.1 GI:14097787
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1560)
AUTHORS Matsuomoto,K. and Nishida,E.
TITLE Tabi protein and DNA coding therefor
JOURNAL Patent: US 6140042-A 1 31-OCT-2000;
FEATURES
Location/Qualifiers
source 1..1560
/organism="unknown"
BASE COUNT 332 a 469 c 480 g 279 t
ORIGIN
Alignment Scores:
Pred. No.: 3.4e-30 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AR116881 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAGCCCGACCTTAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 1398 TCTGACGAGGCGCTCTTCGGCTCCGGCCGCCCACTCGCTCCGGCTGGCGAGACGGT 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517
QY 61 GluGlnSerValValThraPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541
```







Db 1346 CAAAGCCGACCTTAACCTCGAGTCCACCAACGACGAGCAGGAGCTCCAGC 1405  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1406 TCTGACGGAGGCTCTTCCTCCGCGCCGACCACTCGCTCCGCTGGCGAGGACGCT 1465  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1466 CGTGTGAGCCCTATCTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCAAGC 1525  
 QY 61 GluGlnSerValValThrAlaPro 68  
 Db 1526 GAGCAGAGCGTGGTGACGACCG 1549

RESULT 13  
 AR2311195  
 LOCUS AR2311195 1569 bp DNA linear PAT 20-DEC-2002  
 DEFINITION Sequence 10 from patent US 6451617.  
 ACCESSION AR2311195  
 VERSION AR2311195.1 GI:27272083  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1569)  
 AUTHORS Ono,K., Ohtomo,T. and Tsuchiya,M.  
 TITLE Method of screening TGF-beta. inhibitory substances  
 JOURNAL Patent: US 6451617-A 10 17-SEP-2002;  
 FEATURES  
 source Location/Qualifiers  
 1..1569  
 /organism="unknown"  
 BASE COUNT 343 a 466 c 476 g 284 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3 42e-30 Length: 1569  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x AR2311195 (1-1569)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1315 CAAAGCCGACCTTAACCTCGAGTCCACCAACGACGAGCAGGAGCTCCAGC 1374  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1375 TCTGACGGAGGCTCTTCCTCCGCGCCGACCACTCGCTCCGCTGGCGAGGACGCT 1434  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1435 CGTGTGAGCCCTATCTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCAAGC 1494  
 QY 61 GluGlnSerValValThrAlaPro 68  
 Db 1495 GAGCAGAGCGTGGTGACGACCG 1518

RESULT 14  
 AR307976  
 LOCUS AR307976 1569 bp DNA linear PAT 12-JUN-2003  
 DEFINITION Sequence 10 from patent US 6551840.  
 ACCESSION AR307976  
 VERSION AR307976.1 GI:31698733  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1569)  
 AUTHORS Ono,K., Ohtomo,T. and Tsuchiya,M.  
 TITLE Method of screening TGF-beta. inhibiting substances

JOURNAL Patent: US 6551840-A 10 22-APR-2003;  
 FEATURES  
 source Location/Qualifiers  
 1..1569  
 /organism="unknown"  
 BASE COUNT 343 a 466 c 476 g 284 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3 42e-30 Length: 1569  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x AR307976 (1-1569)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 1315 CAAAGCCGACCTTAACCTCGAGTCCACCAACGACGAGCAGGAGCTCCAGC 1374  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 Db 1375 TCTGACGGAGGCTCTTCCTCCGCGCCGACCACTCGCTCCGCTGGCGAGGACGCT 1434  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
 Db 1435 CGTGTGAGCCCTATCTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCAAGC 1494  
 QY 61 GluGlnSerValValThrAlaPro 68  
 Db 1495 GAGCAGAGCGTGGTGACGACCG 1518

RESULT 15  
 HSU49928  
 LOCUS HSU49928 3096 bp mRNA linear PRI 06-APR-1998  
 DEFINITION Homo sapiens TAK1 binding protein (TAK1) mRNA, complete cds.  
 ACCESSION U49928  
 VERSION U49928.1 GI:1401125  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3096)  
 AUTHORS Shibuya,H., Yamaguchi,K., Shirakabe,K., Tonegawa,A., Gotoh,Y.,  
 Ueno,N., Irie,K., Nishida,E. and Matsumoto,K.  
 TITLE TAB1: an activator of the TAK1 MAPKKK in TGF-beta signal transduction  
 JOURNAL Science 272 (5265), 1179-1182 (1996)  
 MEDLINE 96216294  
 PUBMED 8638164

REFERENCE 2 (bases 1 to 3096)  
 AUTHORS Shibuya,H.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-FEB-1996) Hiroshi Shibuya, Faculty of Pharmaceutical Sciences, Hokkaido University, Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan

FEATURES  
 source Location/Qualifiers  
 1..3096  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 gene 1..3096  
 /gene="TAK1"  
 CDS 21..1535  
 /gene="TAK1"  
 /note="activator for TAK1"  
 /codon\_start=1  
 /product="TAK1 binding protein"  
 /protein\_id="AAC12660.1"  
 /db\_xref="GI:1401126"  
 /translation="MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGT"

(5)

ESHPPEDSWLFRSENNCELYGVNGYDGNRVTFVFAORLSAELLIGOLNABEADV  
RRVLLQAFDIVERSFLESIDDALEAKASLQSLPEGVPOHLPPOVQKILERLKTILR  
ETSGGAMAVAVLNNKLYANVTNREALLCKSTVDGLVQTQNLNDHTTENEDELFRLL  
SOLGLDAGIKQVGIICQESTRRIGDYKYKGYTIDIDLLSAKSPHIIAEPEIHGAQ  
PLDGVTFGLVMSGLYKALEAHGPGQANQIAAMIDTEFAKOTSLDVAQAQVVDV  
KRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGEMSQPTSPAPAGGRVYPSV  
PYSSAQSTKTSVTLVMSQCMVNGAHSASTLDEATPTLTNOSPTLTLOSTNHT  
QSSSSSSDGLFRSPAHSLPPGEDGRVEPYVDFAEFYFLMSVDHGEQSVVTAP"

BASE COUNT 642 a 936 c 952 g 566 t  
ORIGIN

## Alignment Scores:

Pred. No.:	6.83e-30	Length:	3096
Score:	359.00	Matches:	68
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-4\_COPY\_437\_504 (1-68) x HSU49928 (1-3096)

QY	1	GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer	20
Db	1329	CAAGCCCGACCTTAACCTTGAGTCCACCAACGACACGACGACGAGCTCCAGC	1388
QY	21	SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly	40
Db	1389	TCTGACGAGGCTCTTCGCTCCCGGCCGCCACTCGCTCCCGCTGGCGAGGACGGT	1448
QY	41	ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly	60
Db	1449	CGTGTTGAGCCCTATGTGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGACCATGGC	1508
QY	61	GluGlnSerValThrAlaPro	68
Db	1509	GAGCAGCGGTGGTACAGCACCG	1532

Search completed: December 4, 2003, 11:40:02  
Job time : 1469.84 secs

*This Page Blank (uspto)*

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:09:50 ; Search time 6.66216 Seconds  
(without alignments)  
479.997 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSPILTQSTNTHQTSSSSS.....AEFYRLWSVDHGQSQVWTP 68

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	359	100.0	504	1	TAB1_HUMAN	Q15750 homo sapien
2	63.5	17.7	243	1	IM17_ARATH	Q9SP35 arabidopsis
3	63.5	17.7	824	1	AD08_HUMAN	P78125 homo sapien
4	62	17.3	541	1	N057_YEAST	P48837 saccharomyc
5	61	17.0	1119	1	ALS3_CANAL	O74623 candida alb
6	61	17.0	1511	1	PRR5_YEAST	P33302 saccharomyc
7	60	16.7	226	1	ORAZ_MOUSE	P54369 mus musculu
8	60	16.7	1260	1	ALSI1_CANAL	P46590 candida alb
9	59	16.4	329	1	CAHX_FLAPR	P46281 flaveria pr
10	59	16.4	330	1	CAH1_FLALI	P46512 flaveria li
11	59	16.4	330	1	CAHX_FLAPR	P46511 flaveria br
12	59	16.4	399	1	PRS3_SHEEP	O97967 ovis aries
13	58.5	16.3	608	1	FLR1_MOUSE	Q98501 mus musculu
14	58.5	16.3	1093	1	AF17_HUMAN	P55198 homo sapien
15	58.5	16.3	1099	1	PLC1_CANAL	O13433 candida alb
16	58.5	16.3	3256	1	K167_HUMAN	P46013 homo sapien
17	58	16.2	142	1	YNEK_BACSU	P45711 bacillus su
18	58	16.2	276	1	CAPB_DROME	P48603 drosophila
19	58	16.2	532	1	SPG7_DICDI	P22698 dictyosteli
20	57.5	16.0	232	1	HXB4_XENLA	P09070 xenopus lae
21	57.5	16.0	396	1	CSA_HUMAN	Q13216 homo sapien
22	57.5	16.0	686	1	VILL_HUMAN	O15195 homo sapien
23	57	15.9	461	1	KEM2_MOUSE	Q9K187 mus musculu
24	57	15.9	937	1	NU98_HUMAN	P52948 homo sapien
25	57	15.9	1377	1	NEO1_RAT	P72603 rattus norv
26	57	15.9	1461	1	NEO1_HUMAN	Q92859 homo sapien
27	56.5	15.7	276	1	PSBS_LYCES	P54773 lycopersico
28	56.5	15.7	283	1	PANC_BACHD	Q9K286 bacillus ha
29	56.5	15.7	343	1	DHSO_BACHD	Q92900 bacillus ha
30	56.5	15.7	457	1	CUC1_ECOLI	P77211 escherichia
31	56.5	15.7	459	1	MCE1_YEAST	Q01159 saccharomyc
32	56.5	15.7	463	1	PLSB_CARTI	O42713 carthamus t
33	56.5	15.7	922	1	NRPI_RAT	Q9QWJ9 rattus norv

34	56.5	15.7	923	1	NRP1_HUMAN	O14786 homo sapien
35	56.5	15.7	1858	1	P3K2_DICDI	P54674 dictyosteli
36	56	15.6	237	1	TG12_HUMAN	Q9GZ12 homo sapien
37	56	15.6	297	1	PRTA_STRGR	P00776 streptomyc
38	56	15.6	315	1	SPY2_HUMAN	O43597 homo sapien
39	56	15.6	325	1	GBLP_SOYEN	Q39836 glycine max
40	56	15.6	405	1	Y574_HUMAN	O60320 homo sapien
41	56	15.6	427	1	TRI6_HUMAN	P08138 homo sapien
42	56	15.6	462	1	KRM2_HUMAN	Q8NCW0 homo sapien
43	56	15.6	509	1	RUNT_DROME	P22814 drosophila
44	56	15.6	639	1	BAR1_SCHCO	Q92275 schizophyll
45	56	15.6	652	1	FXO1_MOUSE	Q9R1E0 mus musculu

ALIGNMENTS

RESULT 1			
ID	TAB1_HUMAN	STANDARD;	PRT; 504 AA.
AC	Q15750;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Mitogen-activated protein kinase kinase 7 interacting protein 1		
DE	(TAK1-binding protein 1).		
GN	MAP3K7ip1 OR TAB1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=96216294; PubMed=8638164;		
RA	Shibuya H., Yamaguchi K., Shirakabe K., Tonegawa A., Gotoh Y.,		
RA	Ueno N., Irie K., Nishida E., Matsumoto K.;		
RT	"TAB1: an activator of the TAK1 MAPKKK in TGF-beta signal		
RT	transduction."		
RL	Science 272:1179-1184 (1996).		
RN	[2]		

[illegible]

Matches 27; Conservative 2; Indels 30; Gaps 3

QY 7 LQSTNTHQSSSSSD-----GGLRSPRAHSLPGEDGRVSPY--VDFAEFVRLWS-- 56  
 Db 161 LQTAGTCGVSDDSLGSLGLRPTAAVFRPRGDSLPSRSTRYVELYVVVDNAEFQMLGSEA 220

QY 57 -VDHGRQSV 65  
 Db 221 AVHRVLEV 230

RESULT 4  
 NUS7 YEAST  
 ID ID NUS7 YEAST STANDARD; PRT; 541 AA.  
 AC P48837;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Nucleoporin NUP57 (Nuclear pore protein NUP57).  
 GN NUP57 OR YGR119C OR G6320.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 NCBI\_TaxID=49332;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95129554; PubMed=7828598;  
 RX Grandi P., Schlaich N., Tekotte H., Hurt E.C.;  
 RA "Functional interaction of Nup57 with a core nucleoporin complex  
 RT consisting of Nup1, Nup49 and a novel protein Nup57p.";  
 RL EMBO J. 14:76-87(1995).  
 RN [2]  
 RP SEQUENCE OF 1-354 FROM N.A.  
 RX STRAIN=S286c / FY1679;  
 RC MEDLINE=97197982; PubMed=9046098;  
 RX van Dyck L., Tettelin H., Purnelle B., Goffeau A.;  
 RA "An 18.3 kb DNA fragment from yeast chromosome VII carries four  
 RT unknown open reading frames, the gene for an Asn synthase, remnants  
 of Ty and three tRNA genes.";  
 RL Yeast 13:171-176(1997).  
 RN [3]  
 RP SEQUENCE OF 243-541 FROM N.A.  
 RX MEDLINE=97061913; PubMed=8905931;  
 RA Hansen M., Albers M., Backes U., Coblentz A., Leuther H., Neu R.,  
 RA Schreier A., Schaefer B., Zimmermann M., Wolf K.;  
 RA "The sequence of a 23.4 kb segment on the right arm of chromosome VII  
 RT from Saccharomyces cerevisiae reveals CLB6, SPT6, RP28A and NUP57  
 RT genes, a Ty3 element and 11 new open reading frames.";  
 RL Yeast 12:1273-1277(1996).  
 CC -!- FUNCTION: PART OF THE NUCLEOPORIN COMPLEX; REQUIRED FOR PROTEIN  
 CC TRANSPORT IN THE NUCLEUS.  
 CC -!- SUBUNIT: INTERACTS WITH NSP1, NUP49 AND NIC96.  
 CC -!- SUBCELLULAR LOCATION: Nuclear pore complex.  
 CC -!- DOMAIN: CONTAINS G-L-F-G REPEATS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X81155; CAA57053.1; -  
 CC EMBL; X83099; CAA58153.1; -  
 CC EMBL; Z72904; CAA97129.1; -  
 CC EMBL; Z72905; CAA97131.1; -  
 CC PIR; S51799; S51799.  
 CC SGD; S0003351; NUP57.  
 CC GO; GO:0005643; C:nuclear pore; IDA.  
 CC GO; GO:0006406; P:mRNA-nucleus export; IMP.  
 CC GO; GO:0006606; P:protein-nucleus import; IMP.  
 CC Nuclear protein; Transport; Coiled coil; Nucleo-

```
FT DOMAIN 76 223 9 X 4 AA REPEATS OF G-L-F-G.
FT DOMAIN 26 31 POLY-ASN.
FT DOMAIN 127 130 POLY-THR.
FT DOMAIN 217 220 POLY-GLY.
FT DOMAIN 258 266 POLY-GLN.
FT DOMAIN 277 280 POLY-GLN.
FT DOMAIN 398 425 COILED COIL (POTENTIAL).
SQ SEQUENCE 541 AA; 57498 MW; 829ADF7B1D7E83C CRC64;

Query Match 17.3%; Score 62; DB 1; Length 541;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 12; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 8 QSTNTHQSSSSSDGGLFRSRPA 31
:|||||:|||||:|||||:
DB 60 QATNFGSNQSSSTGGGLFGNKPA 83

RESULT 5
ALS3-CANAL STANDARD; PRT; 1119 AA.
AC 074623;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Agglutinin-like protein 3 precursor.
GN ALS3.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=11161;
RX MEDLINE=98309840; PubMed=9644209;
RA Hoyer L.L., Payne T.L., Bell M., Myers A.M., Scherer S.;
RT "Candida albicans ALS3 and insights into the nature of the ALS gene family.";
RL Curr. Genet. 33:451-459(1998).
CC -!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
CC -!- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U87956; AAC39486.1; -.
KW Cell adhesion; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 1119
FT DOMAIN 433 792 AGGLUTININ-LIKE PROTEIN 3.
FT REPEAT 433 468 10 X 36 AA TANDEM REPEATS.
FT REPEAT 469 504 1-1.
FT REPEAT 505 540 1-2.
FT REPEAT 541 576 1-3.
FT REPEAT 577 612 1-4.
FT REPEAT 613 648 1-5.
FT REPEAT 649 684 1-6.
FT REPEAT 685 720 1-7.
FT REPEAT 721 756 1-8.
FT REPEAT 757 792 1-9.
FT REPEAT 793 829 1-10.
FT DOMAIN 399 404 POLY-THR.
FT DOMAIN 450 455 POLY-THR.
FT DOMAIN 557 563 POLY-THR.
FT DOMAIN 593 597 POLY-THR.
FT DOMAIN 630 635 POLY-THR.
FT DOMAIN 666 671 POLY-THR.
FT DOMAIN 702 707 POLY-THR.
FT DOMAIN 738 743 POLY-THR.
```

```
FT DOMAIN 774 777 POLY-THR.
FT DOMAIN 1044 1047 POLY-THR.
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 687 687 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 723 723 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 759 759 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 845 845 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 987 987 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1050 1050 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1119 AA; 119927 MW; 6A3FB3FC8C879A71 CRC64;

Query Match 17.0%; Score 61; DB 1; Length 1119;
Best Local Similarity 23.8%; Pred. No. 44;
Matches 24; Conservative 13; Mismatches 28; Indels 36; Gaps 4;

QY 1 QSPTLTLOSTNTHQSSSSS-----SDGGLFRSRPAHSL----- 34
:|||||:|||||:|||||:
DB 720 EPNHVTITTEYWSQSYATTTTAPPGEDTDLVLRPPNHTVTTEYWSQSYATTTTII 779
:|||||:|||||:|||||:
QY 35 -PREGEDRV-----EPYVDFAEFYRLWSVDHGQSVVTAP 68
|||||:|||||:|||||:
DB 780 APGETDTVLIRPPNPTVTTEY---WSQSYTTATTVTAP 817

RESULT 6
PDR5 YEAST STANDARD; PRT; 1511 AA.
AC P33302;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Suppressor of toxicity of sporidesmin.
GN PDR5 OR STS1 OR YDR1 OR LEM1 OR YOR153W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB320;
RX MEDLINE=94140838; PubMed=9307980;
RA Bessinger P.H., Kuchler K.;
RT "Molecular cloning and expression of the Saccharomyces cerevisiae
RT STS1 gene product. A yeast ABC transporter conferring mycotoxin
RT resistance.";
RL J. Biol. Chem. 269:4180-4186 (1994).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=S288C;
RX MEDLINE=94124579; PubMed=8294477;
RA Balzi E., Wang M., Leterme S., van Dyck L., Goffeau A.;
RT "PDR5, a novel yeast multidrug resistance conferring transporter
RT controlled by the transcription regulator PDR1.";
RL J. Biol. Chem. 269:2206-2214 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AH22;
RX MEDLINE=95188264; PubMed=7882421;
RA Hirata D., Yano K., Miyahara K., Miyakawa T.;
RT "Saccharomyces cerevisiae YDR1, which encodes a member of the ATP-
RT binding cassette (ABC) superfamily, is required for multidrug
RT resistance.";
RL Curr. Genet. 26:285-294 (1994).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1678;
RA Ayadi A., Bordonne R., Camasses A., Madania A., Poch O.,
RA Tarasov I.A., Winsor B., Martin R.P.;
RT "Analysis of a 35600 bp region on the right arm of Saccharomycetes
```



```

cervisiae chromosome XV."
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSPORTER, WHOSE PHYSIOLOGICAL FUNCTION IS NOT YET
CC ESTABLISHED. CONFERS RESISTANCE TO THE CHEMICALS CYCLOHEXIMIDE
CC AND SULFOMETHURON METHYL. EXHIBITS NUCLEOSIDE TRIPHOSPHATASE
CC ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. PDR5 SUBFAMILY.
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 61.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X74113; CAA52212.1; -
CC EMBL; L19922; AAB53769.1; -
CC EMBL; D28548; BAA05547.1; ALT_INIT.
CC EMBL; U55020; AAC49639.1; -
CC EMBL; Z75061; CAA99359.1; -
CC PIR; A53151; A53151.
CC SGD; S0005679; PDR5.
CC GO; GO:0005886; C:plasma membrane; IDA.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR005285; PDR.
CC Pfam; PF00005; ABC_tran; 2.
CC ProDom; PD000006; ABC_transporter; 2.
CC SMART; SM00382; AAA; 2.
CC TIGRfams; TIGR00956; 3a01205; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 2.
CC ATP-binding; Transmembrane; Glycoprotein; Transport.
CC DOMAIN 1 517 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 518 542 POTENTIAL.
CC TRANSMEM 559 579 POTENTIAL.
CC TRANSMEM 612 628 POTENTIAL.
CC TRANSMEM 632 650 POTENTIAL.
CC TRANSMEM 666 685 POTENTIAL.
CC TRANSMEM 775 793 POTENTIAL.
CC DOMAIN 794 1237 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 1238 1260 POTENTIAL.
CC TRANSMEM 1292 1313 POTENTIAL.
CC TRANSMEM 1325 1349 POTENTIAL.
CC TRANSMEM 1355 1379 POTENTIAL.
CC TRANSMEM 1389 1407 POTENTIAL.
CC TRANSMEM 1477 1499 POTENTIAL.
CC DOMAIN 1500 1511 CYTOPLASMIC (POTENTIAL).
CC NP_BIND 905 912 ATP (POTENTIAL).
CC DOMAIN 784 787 POLY-PHE.
CC CONFLICT 171 171 N -> L (IN REF. 3).
CC CONFLICT 190 190 V -> I (IN REF. 3).
CC CONFLICT 214 214 D -> T (IN REF. 3).
CC CONFLICT 308 308 G -> V (IN REF. 3).
CC CONFLICT 340 345 MISSING (IN REF. 3).
CC CONFLICT 476 476 R -> H (IN REF. 3).
CC CONFLICT 648 648 MISSING (IN REF. 3).
CC CONFLICT 770 770 D -> H (IN REF. 3).
CC SEQUENCE 1511 AA; 170437 MW; 4540DC0BF0474BA CRC64;

Query Match 17.0%; Score 61; DB 1; Length 1511;
Best Local Similarity 38.8; Pred. No. 64;
Matches 19; Conservative 6; Mismatches 16; Indels 8; Gaps 3;

QY 4 TLTLQSTNTHQTSSSSSDG-GLFRS-----RPAHSLP--PGDGRVPEP 44
DB 49 TLTAQSMQNSTQAPNKSQAQSFSSGVEGVNPIFSDPEAPGYDPKLD 97

```

```

RESULT 7
OAZ_MOUSE STANDARD; PRT; 226 AA.
ID OAZ_MOUSE
AC P54369; O08610;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ornithine decarboxylase antizyme (ODC-Az).
GN OAZ1 OR OAZ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9808944; PubMed=9428668;
RA Nilsson J., Koskinen S., Persson K., Grahn B., Holm I.;
RT "Polyamines regulate both transcription and translation of the gene
RT encoding ornithine decarboxylase antizyme in mouse.";
RT Eur. J. Biochem. 250:223-231(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Svj;
RA Kankare K., Uusi-Oukari M., Janne O.A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BINDS TO, AND DESTABILIZES, ORNITHINE DECARBOXYLASE
CC WHICH IS THEN DEGRADED. ALSO INHIBITS CELLULAR UPTAKE OF
CC POLYAMINES BY INACTIVATING THE POLYAMINE UPTAKE TRANSPORTER.
CC -!- MISCELLANEOUS: A RIBOSOMAL FRAMESHIFT OCCURS BETWEEN THE CODONS
CC FOR SER-67 AND ASP-68. AN AUTOREGULATORY MECHANISM ENABLES
CC MODULATION OF FRAMESHIFTING ACCORDING TO THE CELLULAR
CC CONCENTRATION OF POLYAMINES.
CC -!- SIMILARITY: BELONGS TO THE ODC ANTIZYME FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U52822; AAB96329.1; -
CC EMBL; U52823; AAB96330.1; -
CC EMBL; U84291; AAC53307.1; -
CC MGD; MGI:109433; Oaz1.
CC InterPro; IPR002993; ODC_AZ.
CC Pfam; PF02100; ODC_AZ; 1.
CC ProDom; PD007483; ODC_AZ; 1.
CC PROSITE; PS01337; ODC_AZ; 1.
CC Ribosomal frameshift.
CC INIT MET 0 0 BY SIMILARITY.
CC CONFLICT 68 68 D -> C (IN REF. 1: AAB96330).
CC SEQUENCE 226 AA; 25002 MW; C58D8D6B730318F6 CRC64;

Query Match 16.7%; Score 60; DB 1; Length 226;
Best Local Similarity 31.7%; Pred. No. 8.4;
Matches 19; Conservative 8; Mismatches 23; Indels 10; Gaps 2;

QY 1 QSTT-----LTLOSTNTHQTSSSSSDGGLFRSRAHSLPPGDGRVPEYVDPAE 50
DB 104 EEPSTNDKTRVLSTQTLTEAKQVTRAVWSGGGLYELPAGLPPEGSKDSFAALLEFAE 163

RESULT 8
ALSL1_CANAL STANDARD; PRT; 1260 AA.
ID ALSL1_CANAL
AC P46590;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Agglutinin-like protein 1 precursor.
GN ALS1.

```

OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 11651 / B792;  
 RA MEDLINE=95272392; PubMed=7752895;  
 RX Hoyer L.L., Scherer S., Shatzman A.R., Livi G.P.;  
 RT "Candida albicans ALS1: domains related to a Saccharomyces cerevisiae  
 RT sexual agglutinin separated by a repeating motif.";  
 RL Mol. Microbiol. 15:39-54(1995).  
 CC -!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.  
 CC -!- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).  
 CC -!- SIMILARITY: TO YEAST SAG1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; L25902; AAC41649.2;  
 CC Cell adhesion; Glycoprotein; Repeat; Signal.  
 KW SIGNAL  
 FT CHAIN 1 17  
 FT DOMAIN 18 1260 AGGLUTININ-LIKE PROTEIN 1.  
 FT REPEAT 433 792 10 X 36 AA TANDEM REPEATS.  
 FT REPEAT 433 468 1-1.  
 FT REPEAT 469 504 1-2.  
 FT REPEAT 505 540 1-3.  
 FT REPEAT 541 576 1-4.  
 FT REPEAT 577 612 1-5.  
 FT REPEAT 613 648 1-6.  
 FT REPEAT 649 684 1-7.  
 FT REPEAT 685 720 1-8.  
 FT REPEAT 721 756 1-9.  
 FT REPEAT 757 792 1-10.  
 FT DOMAIN 983 1152 2 X 26 AA APPROXIMATE REPEATS.  
 FT REPEAT 983 1043 2-1.  
 FT REPEAT 1092 1152 2-2.  
 FT DOMAIN 399 404 POLY-THR.  
 FT DOMAIN 408 418 POLY-THR.  
 FT DOMAIN 450 455 POLY-THR.  
 FT DOMAIN 486 491 POLY-THR.  
 FT DOMAIN 522 527 POLY-THR.  
 FT DOMAIN 558 563 POLY-THR.  
 FT DOMAIN 594 599 POLY-THR.  
 FT DOMAIN 630 635 POLY-THR.  
 FT DOMAIN 666 671 POLY-THR.  
 FT DOMAIN 702 707 POLY-THR.  
 FT DOMAIN 738 743 POLY-THR.  
 FT DOMAIN 774 779 POLY-THR.  
 FT DOMAIN 874 877 POLY-SER.  
 FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 687 687 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 723 723 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 820 820 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 886 886 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 918 918 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 973 973 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1068 1068 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 1260 AA; 132641 MW; 763D1063A2354C24 CRC64;  
 Query Match 16.7%; Score 60; DB 1; Length 1260;  
 Best Local Similarity 24.2%; Pred. No. 67;  
 Matches 23; Conservative 9; Mismatches 35; Indels 28; Gaps 3;  
 2 SPTLTL-----QSTNTHQSSSSSDGGLFRSRAHSL-----PP 36

Db 651 NPVTITTEYSQSYATTITITAPPGBTDTVLIREPNNHTVTTEYSQSYATTITVTAPP 710  
 QY 37 GEDGRV---EPYVDFAEFRLWSVDHGEQSVVTAP 68  
 Db 711 GETDTVLIREPPNHTVTTEYSQSYATTITVTAP 745  
 RESULT 9  
 CAHX FLAPR STANDARD; PRT; 329 AA.  
 ID CAHX FLAPR STANDARD; PRT; 329 AA.  
 AC P46281;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Carbonic anhydrase (EC 4.2.1.1) (Carbonate dehydratase).  
 OS Flaveria pringlei.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Asteridae; Campanulids; Asterales; Asteraceae; Asteroideae; Tageteae;  
 CC Flaveria.  
 OX NCBI\_TaxID=4226;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Leaf;  
 RX MEDLINE=96046753; PubMed=7579185;  
 RA Ludwig M., Burnell J.N.;  
 RT "Molecular comparison of carbonic anhydrase from Flaveria species  
 RT demonstrating different photosynthetic pathways.";  
 RL Plant Mol. Biol. 29:353-365(1995).  
 CC -!- FUNCTION: Reversible hydration of carbon dioxide.  
 CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.  
 CC -!- SUBUNIT: Homohexamer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (potential).  
 CC -!- DOMAIN: POSSESSES A TRANSIT-LIKE PEPTIDE, BUT IT IS PROPOSED THAT  
 CC THIS PEPTIDE IS NOT REMOVED AND THAT THEREFORE THE ENZYME STAYS  
 CC IN THE CYTOPLASM INSTEAD OF GOING TO THE CHLOROPLAST  
 CC (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE PLANT AND PROKARYOTIC CARBONIC  
 CC ANHYDRASE FAMILY.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; U19737; AAA86992.1; -  
 CC FIR; S61884; S61884.  
 CC InterPro; IPR001765; Prok\_Coanhd.  
 CC Pfam; PF00484; Pro\_CA; 1.  
 CC PROSITE; PS00704; PROK\_CO2\_ANHYDRASE\_1; 1.  
 CC PROSITE; PS00705; PROK\_CO2\_ANHYDRASE\_2; 1.  
 CC Lyase; Zinc. 1 108 CHLOROPLAST TRANSIT PEPTIDE-LIKE.  
 FT DOMAIN 23 26 POLY-SER.  
 FT DOMAIN 40 46 POLY-SER.  
 SQ SEQUENCE 329 AA; 35486 MW; B18E656B1E84C34B CRC64;  
 Query Match 16.4%; Score 59; DB 1; Length 329;  
 Best Local Similarity 27.0%; Pred. No. 17;  
 Matches 20; Conservative 13; Mismatches 33; Indels 8; Gaps 2;  
 1 QSPITLITQSTNTHQSSSSSDGGLFRSRAHSLPP-----GEDGRVEPYVDFAEFVR 53  
 Db 28 RSGVLSARFTCNSSSSSSSSTPPSLIRNEPFAAPAIITPNWTDGN-ESYEIDAALK 86  
 QY 54 LWSVDHGEQSVVTA 67  
 Db 87 KMLIEKGELEFVAA 100

[illegible]

RC TISSUE=Pituitary;  
RX MEDLINE=99367324; PubMed=10425452;  
RA Whitley J.C., Moore C., Giraud A.S., Shulkes A.;  
RT "Molecular cloning, genomic organization and selective expression of  
RT bombesin receptor subtype 3 in the sheep hypothalamus and  
RT pituitary";  
RL J. Mol. Endocrinol. 23:107-116(1999).  
CC -!- FUNCTION: ROLE IN SPERM CELL DIVISION, MATURATION, OR FUNCTION.  
CC THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS.  
CC THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER  
CC SYSTEM (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF108210; AAD19642.1; -  
CC EMBL; AF108209; AAD19639.1; -  
CC EMBL; AF108207; AAD19639.1; JOINED.  
CC EMBL; AF108208; AAD19639.1; JOINED.  
CC InterPro; IPR000276; GPCR\_Rhodops.  
CC Pfam; PF00001; 7tm.1; 1.  
CC PRINTS; PR00237; GPCRHHODOPS.  
CC PROSITE; PS00237; G-PROTEIN RECF\_F1\_1; 1.  
CC PROSITE; PS00262; G-PROTEIN RECF\_F2\_1; 1.  
CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
KW Lipoprotein; Palmitate.  
FT DOMAIN 1 41  
FT TRANSMEM 42 63  
FT DOMAIN 64 82  
FT TRANSMEM 83 103  
FT DOMAIN 104 121  
FT TRANSMEM 122 143  
FT DOMAIN 144 163  
FT TRANSMEM 164 184  
FT DOMAIN 185 220  
FT TRANSMEM 221 241  
FT DOMAIN 242 272  
FT TRANSMEM 273 293  
FT DOMAIN 294 313  
FT TRANSMEM 314 333  
FT DOMAIN 334 399  
FT CARBOHYD 10 10  
FT CARBOHYD 18 18  
FT CARBOHYD 29 29  
FT DISULFID 120 203  
SQ SEQUENCE 399 AA; 44373 MW; B48DD27197AED2EB CRC64;  
Query Match 16.4%; Score 59; DB 1; Length 399;  
Best Local Similarity 34.7%; Pred. No. 22;  
Matches 17; Conservative 4; Mismatches 28; Indels 0; Gaps 0;  
QY 1 QSPFTLTQSTNTHQSSSSSDGGLFRSPAHSLPPGDDGRVEPYVDFA 49  
DB 7 QSPNQTLISTNDIESSSVVPDSTNKRRTGDSFGIEALCAIYIYA 55  
RESULT 13  
PRLR MOUSE STANDARD; PRT; 608 AA.  
AC Q08501; P15212; P15213; Q62099;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Prolactin receptor precursor (PRL-R).  
DN PRLR.  
OS Mus musculus (Mouse).

OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM PRL-R3).  
RC STRAIN=C3H; TISSUE=Mammary gland;  
RX MEDLINE=94085788; PubMed=8262385;  
RA Moore R.C., Oka T.;  
RT "Cloning and sequencing of the cDNA encoding the murine mammary gland  
RT long-form prolactin receptor.";  
RL Gene 134:263-265(1993).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM PRL-R3).  
RC STRAIN=Swiss Webster; TISSUE=Liver;  
RX MEDLINE=93307149; PubMed=8319571;  
RA Clarke D.L., Linzer D.I.H.;  
RT "Changes in prolactin receptor expression during pregnancy in the  
RT mouse ovary.";  
RL Endocrinology 133:224-232(1993).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM PRL-R3).  
RC STRAIN=Swiss Webster; TISSUE=Liver;  
RX MEDLINE=89261824; PubMed=2725531;  
RA Davis J.A., Linzer D.I.H.;  
RT "Expression of multiple forms of the prolactin receptor in mouse  
RT liver.";  
RL Mol. Endocrinol. 3:674-680(1989).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM PRL-R3).  
RC STRAIN=BALB/c; TISSUE=Mammary gland;  
RX MEDLINE=89261824; PubMed=2725531;  
RA Ebery M., Pezet A., Nandi S., Kelly P.A.;  
RT Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORMS PRL-R2 AND PRL-R1).  
RC STRAIN=Swiss Webster; TISSUE=Liver;  
RX MEDLINE=89261824; PubMed=2725531;  
RA Davis J.A., Linzer D.I.H.;  
RT "Expression of multiple forms of the prolactin receptor in mouse  
RT liver.";  
RL Mol. Endocrinol. 3:674-680(1989).  
RN [6]  
RP FUNCTION: THIS IS A RECEPTOR FOR THE ANTERIOR PITUITARY HORMONE  
RC PROLACTIN, AS WELL AS PLACENTAL LACTOGEN I AND II.  
RX SUBCELLULAR LOCATION: Type I membrane protein.  
RN [7]  
RP ALTERNATIVE PRODUCTS:  
RC Event=Alternative splicing; Named isoforms=3;  
RX Name=PRL-R3;  
RN IsoId=Q08501-1; Sequence=Displayed;  
RX Name=PRL-R1;  
RN IsoId=Q08501-2; Sequence=VSP\_001723, VSP\_001724;  
RX Name=PRL-R2;  
RN IsoId=Q08501-3; Sequence=VSP\_001721, VSP\_001722;  
RX Name=PRL-R3;  
RN IsoId=Q08501-4; Sequence=VSP\_001725, VSP\_001726;  
RX Name=PRL-R4;  
RN IsoId=Q08501-5; Sequence=VSP\_001727, VSP\_001728;  
RX Name=PRL-R5;  
RN IsoId=Q08501-6; Sequence=VSP\_001729, VSP\_001730;  
RX Name=PRL-R6;  
RN IsoId=Q08501-7; Sequence=VSP\_001731, VSP\_001732;  
RX Name=PRL-R7;  
RN IsoId=Q08501-8; Sequence=VSP\_001733, VSP\_001734;  
RX Name=PRL-R8;  
RN IsoId=Q08501-9; Sequence=VSP\_001735, VSP\_001736;  
RX Name=PRL-R9;  
RN IsoId=Q08501-10; Sequence=VSP\_001737, VSP\_001738;  
RX Name=PRL-R10;  
RN IsoId=Q08501-11; Sequence=VSP\_001739, VSP\_001740;  
RX Name=PRL-R11;  
RN IsoId=Q08501-12; Sequence=VSP\_001741, VSP\_001742;  
RX Name=PRL-R12;  
RN IsoId=Q08501-13; Sequence=VSP\_001743, VSP\_001744;  
RX Name=PRL-R13;  
RN IsoId=Q08501-14; Sequence=VSP\_001745, VSP\_001746;  
RX Name=PRL-R14;  
RN IsoId=Q08501-15; Sequence=VSP\_001747, VSP\_001748;  
RX Name=PRL-R15;  
RN IsoId=Q08501-16; Sequence=VSP\_001749, VSP\_001750;  
RX Name=PRL-R16;  
RN IsoId=Q08501-17; Sequence=VSP\_001751, VSP\_001752;  
RX Name=PRL-R17;  
RN IsoId=Q08501-18; Sequence=VSP\_001753, VSP\_001754;  
RX Name=PRL-R18;  
RN IsoId=Q08501-19; Sequence=VSP\_001755, VSP\_001756;  
RX Name=PRL-R19;  
RN IsoId=Q08501-20; Sequence=VSP\_001757, VSP\_001758;  
RX Name=PRL-R20;  
RN IsoId=Q08501-21; Sequence=VSP\_001759, VSP\_001760;  
RX Name=PRL-R21;  
RN IsoId=Q08501-22; Sequence=VSP\_001761, VSP\_001762;  
RX Name=PRL-R22;  
RN IsoId=Q08501-23; Sequence=VSP\_001763, VSP\_001764;  
RX Name=PRL-R23;  
RN IsoId=Q08501-24; Sequence=VSP\_001765, VSP\_001766;  
RX Name=PRL-R24;  
RN IsoId=Q08501-25; Sequence=VSP\_001767, VSP\_001768;  
RX Name=PRL-R25;  
RN IsoId=Q08501-26; Sequence=VSP\_001769, VSP\_001770;  
RX Name=PRL-R26;  
RN IsoId=Q08501-27; Sequence=VSP\_001771, VSP\_001772;  
RX Name=PRL-R27;  
RN IsoId=Q08501-28; Sequence=VSP\_001773, VSP\_001774;  
RX Name=PRL-R28;  
RN IsoId=Q08501-29; Sequence=VSP\_001775, VSP\_001776;  
RX Name=PRL-R29;  
RN IsoId=Q08501-30; Sequence=VSP\_001777, VSP\_001778;  
RX Name=PRL-R30;  
RN IsoId=Q08501-31; Sequence=VSP\_001779, VSP\_001780;  
RX Name=PRL-R31;  
RN IsoId=Q08501-32; Sequence=VSP\_001781, VSP\_001782;  
RX Name=PRL-R32;  
RN IsoId=Q08501-33; Sequence=VSP\_001783, VSP\_001784;  
RX Name=PRL-R33;  
RN IsoId=Q08501-34; Sequence=VSP\_001785, VSP\_001786;  
RX Name=PRL-R34;  
RN IsoId=Q08501-35; Sequence=VSP\_001787, VSP\_001788;  
RX Name=PRL-R35;  
RN IsoId=Q08501-36; Sequence=VSP\_001789, VSP\_001790;  
RX Name=PRL-R36;  
RN IsoId=Q08501-37; Sequence=VSP\_001791, VSP\_001792;  
RX Name=PRL-R37;  
RN IsoId=Q08501-38; Sequence=VSP\_001793, VSP\_001794;  
RX Name=PRL-R38;  
RN IsoId=Q08501-39; Sequence=VSP\_001795, VSP\_001796;  
RX Name=PRL-R39;  
RN IsoId=Q08501-40; Sequence=VSP\_001797, VSP\_001798;  
RX Name=PRL-R40;  
RN IsoId=Q08501-41; Sequence=VSP\_001799, VSP\_001800;  
RX Name=PRL-R41;  
RN IsoId=Q08501-42; Sequence=VSP\_001801, VSP\_001802;  
RX Name=PRL-R42;  
RN IsoId=Q08501-43; Sequence=VSP\_001803, VSP\_001804;  
RX Name=PRL-R43;  
RN IsoId=Q08501-44; Sequence=VSP\_001805, VSP\_001806;  
RX Name=PRL-R44;  
RN IsoId=Q08501-45; Sequence=VSP\_001807, VSP\_001808;  
RX Name=PRL-R45;  
RN IsoId=Q08501-46; Sequence=VSP\_001809, VSP\_001810;  
RX Name=PRL-R46;  
RN IsoId=Q08501-47; Sequence=VSP\_001811, VSP\_001812;  
RX Name=PRL-R47;  
RN IsoId=Q08501-48; Sequence=VSP\_001813, VSP\_001814;  
RX Name=PRL-R48;  
RN IsoId=Q08501-49; Sequence=VSP\_001815, VSP\_001816;  
RX Name=PRL-R49;  
RN IsoId=Q08501-50; Sequence=VSP\_001817, VSP\_001818;  
RX Name=PRL-R50;  
RN IsoId=Q08501-51; Sequence=VSP\_001819, VSP\_001820;  
RX Name=PRL-R51;  
RN IsoId=Q08501-52; Sequence=VSP\_001821, VSP\_001822;  
RX Name=PRL-R52;  
RN IsoId=Q08501-53; Sequence=VSP\_001823, VSP\_001824;  
RX Name=PRL-R53;  
RN IsoId=Q08501-54; Sequence=VSP\_001825, VSP\_001826;  
RX Name=PRL-R54;  
RN IsoId=Q08501-55; Sequence=VSP\_001827, VSP\_001828;  
RX Name=PRL-R55;  
RN IsoId=Q08501-56; Sequence=VSP\_001829, VSP\_001830;  
RX Name=PRL-R56;  
RN IsoId=Q08501-57; Sequence=VSP\_001831, VSP\_001832;  
RX Name=PRL-R57;  
RN IsoId=Q08501-58; Sequence=VSP\_001833, VSP\_001834;  
RX Name=PRL-R58;  
RN IsoId=Q08501-59; Sequence=VSP\_001835, VSP\_001836;  
RX Name=PRL-R59;  
RN IsoId=Q08501-60; Sequence=VSP\_001837, VSP\_001838;  
RX Name=PRL-R60;  
RN IsoId=Q08501-61; Sequence=VSP\_001839, VSP\_001840;  
RX Name=PRL-R61;  
RN IsoId=Q08501-62; Sequence=VSP\_001841, VSP\_001842;  
RX Name=PRL-R62;  
RN IsoId=Q08501-63; Sequence=VSP\_001843, VSP\_001844;  
RX Name=PRL-R63;  
RN IsoId=Q08501-64; Sequence=VSP\_001845, VSP\_001846;  
RX Name=PRL-R64;  
RN IsoId=Q08501-65; Sequence=VSP\_001847, VSP\_001848;  
RX Name=PRL-R65;  
RN IsoId=Q08501-66; Sequence=VSP\_001849, VSP\_001850;  
RX Name=PRL-R66;  
RN IsoId=Q08501-67; Sequence=VSP\_001851, VSP\_001852;  
RX Name=PRL-R67;  
RN IsoId=Q08501-68; Sequence=VSP\_001853, VSP\_001854;  
RX Name=PRL-R68;  
RN IsoId=Q08501-69; Sequence=VSP\_001855, VSP\_001856;  
RX Name=PRL-R69;  
RN IsoId=Q08501-70; Sequence=VSP\_001857, VSP\_001858;  
RX Name=PRL-R70;  
RN IsoId=Q08501-71; Sequence=VSP\_001859, VSP\_001860;  
RX Name=PRL-R71;  
RN IsoId=Q08501-72; Sequence=VSP\_001861, VSP\_001862;  
RX Name=PRL-R72;  
RN IsoId=Q08501-73; Sequence=VSP\_001863, VSP\_001864;  
RX Name=PRL-R73;  
RN IsoId=Q08501-74; Sequence=VSP\_001865, VSP\_001866;  
RX Name=PRL-R74;  
RN IsoId=Q08501-75; Sequence=VSP\_001867, VSP\_001868;  
RX Name=PRL-R75;  
RN IsoId=Q08501-76; Sequence=VSP\_001869, VSP\_001870;  
RX Name=PRL-R76;  
RN IsoId=Q08501-77; Sequence=VSP\_001871, VSP\_001872;  
RX Name=PRL-R77;  
RN IsoId=Q08501-78; Sequence=VSP\_001873, VSP\_001874;  
RX Name=PRL-R78;  
RN IsoId=Q08501-79; Sequence=VSP\_001875, VSP\_001876;  
RX Name=PRL-R79;  
RN IsoId=Q08501-80; Sequence=VSP\_001877, VSP\_001878;  
RX Name=PRL-R80;  
RN IsoId=Q08501-81; Sequence=VSP\_001879, VSP\_001880;  
RX Name=PRL-R81;  
RN IsoId=Q08501-82; Sequence=VSP\_001881, VSP\_001882;  
RX Name=PRL-R82;  
RN IsoId=Q08501-83; Sequence=VSP\_001883, VSP\_001884;  
RX Name=PRL-R83;  
RN IsoId=Q08501-84; Sequence=VSP\_001885, VSP\_001886;  
RX Name=PRL-R84;  
RN IsoId=Q08501-85; Sequence=VSP\_001887, VSP\_001888;  
RX Name=PRL-R85;  
RN IsoId=Q08501-86; Sequence=VSP\_001889, VSP\_001890;  
RX Name=PRL-R86;  
RN IsoId=Q08501-87; Sequence=VSP\_001891, VSP\_001892;  
RX Name=PRL-R87;  
RN IsoId=Q08501-88; Sequence=VSP\_001893, VSP\_001894;  
RX Name=PRL-R88;  
RN IsoId=Q08501-89; Sequence=VSP\_001895, VSP\_001896;  
RX Name=PRL-R89;  
RN IsoId=Q08501-90; Sequence=VSP\_001897, VSP\_001898;  
RX Name=PRL-R90;  
RN IsoId=Q08501-91; Sequence=VSP\_001899, VSP\_001900;  
RX Name=PRL-R91;  
RN IsoId=Q08501-92; Sequence=VSP\_001901, VSP\_001902;  
RX Name=PRL-R92;  
RN IsoId=Q08501-93; Sequence=VSP\_001903, VSP\_001904;  
RX Name=PRL-R93;  
RN IsoId=Q08501-94; Sequence=VSP\_001905, VSP\_001906;  
RX Name=PRL-R94;  
RN IsoId=Q08501-95; Sequence=VSP\_001907, VSP\_001908;  
RX Name=PRL-R95;  
RN IsoId=Q08501-96; Sequence=VSP\_001909, VSP\_001910;  
RX Name=PRL-R96;  
RN IsoId=Q08501-97; Sequence=VSP\_001911, VSP\_001912;  
RX Name=PRL-R97;  
RN IsoId=Q08501-98; Sequence=VSP\_001913, VSP\_001914;  
RX Name=PRL-R98;  
RN IsoId=Q08501-99; Sequence=VSP\_001915, VSP\_001916;  
RX Name=PRL-R99;  
RN IsoId=Q08501-100; Sequence=VSP\_001917, VSP\_001918;  
RX Name=PRL-R100;  
RN IsoId=Q08501-101; Sequence=VSP\_001919, VSP\_001920;  
RX Name=PRL-R101;  
RN IsoId=Q08501-102; Sequence=VSP\_001921, VSP\_001922;  
RX Name=PRL-R102;  
RN IsoId=Q08501-103; Sequence=VSP\_001923, VSP\_001924;  
RX Name=PRL-R103;  
RN IsoId=Q08501-104; Sequence=VSP\_001925, VSP\_001926;  
RX Name=PRL-R104;  
RN IsoId=Q08501-105; Sequence=VSP\_001927, VSP\_001928;  
RX Name=PRL-R105;  
RN IsoId=Q08501-106; Sequence=VSP\_001929, VSP\_001930;  
RX Name=PRL-R106;  
RN IsoId=Q08501-107; Sequence=VSP\_001931, VSP\_001932;  
RX Name=PRL-R107;  
RN IsoId=Q08501-108; Sequence=VSP\_001933, VSP\_001934;  
RX Name=PRL-R108;  
RN IsoId=Q08501-109; Sequence=VSP\_001935, VSP\_001936;  
RX Name=PRL-R109;  
RN IsoId=Q08501-110; Sequence=VSP\_001937, VSP\_001938;  
RX Name=PRL-R110;  
RN IsoId=Q08501-111; Sequence=VSP\_001939, VSP\_001940;  
RX Name=PRL-R111;  
RN IsoId=Q08501-112; Sequence=VSP\_001941, VSP\_001942;  
RX Name=PRL-R112;  
RN IsoId=Q08501-113; Sequence=VSP\_001943, VSP\_001944;  
RX Name=PRL-R113;  
RN IsoId=Q08501-114; Sequence=VSP\_001945, VSP\_001946;  
RX Name=PRL-R114;  
RN IsoId=Q08501-115; Sequence=VSP\_001947, VSP\_001948;  
RX Name=PRL-R115;  
RN IsoId=Q08501-116; Sequence=VSP\_001949, VSP\_001950;  
RX Name=PRL-R116;  
RN IsoId=Q08501-117; Sequence=VSP\_001951, VSP\_001952;  
RX Name=PRL-R117;  
RN IsoId=Q08501-118; Sequence=VSP\_001953, VSP\_001954;  
RX Name=PRL-R118;  
RN IsoId=Q08501-119; Sequence=VSP\_001955, VSP\_001956;  
RX Name=PRL-R119;  
RN IsoId=Q08501-120; Sequence=VSP\_001957, VSP\_001958;  
RX Name=PRL-R120;  
RN IsoId=Q08501-121; Sequence=VSP\_001959, VSP\_001960;  
RX Name=PRL-R121;  
RN IsoId=Q08501-122; Sequence=VSP\_001961, VSP\_001962;  
RX Name=PRL-R122;  
RN IsoId=Q08501-123; Sequence=VSP\_001963, VSP\_001964;  
RX Name=PRL-R123;  
RN IsoId=Q08501-124; Sequence=VSP\_001965, VSP\_001966;  
RX Name=PRL-R124;  
RN IsoId=Q08501-125; Sequence=VSP\_001967, VSP\_001968;  
RX Name=PRL-R125;  
RN IsoId=Q08501-126; Sequence=VSP\_001969, VSP\_001970;  
RX Name=PRL-R126;  
RN IsoId=Q08501-127; Sequence=VSP\_001971, VSP\_001972;  
RX Name=PRL-R127;  
RN IsoId=Q08501-128; Sequence=VSP\_001973, VSP\_001974;  
RX Name=PRL-R128;  
RN IsoId=Q08501-129; Sequence=VSP\_001975, VSP\_001976;  
RX Name=PRL-R129;  
RN IsoId=Q08501-130; Sequence=VSP\_001977, VSP\_001978;  
RX Name=PRL-R130;  
RN IsoId=Q08501-131; Sequence=VSP\_001979, VSP\_001980;  
RX Name=PRL-R131;  
RN IsoId=Q08501-132; Sequence=VSP\_001981, VSP\_001982;  
RX Name=PRL-R132;  
RN IsoId=Q08501-133; Sequence=VSP\_001983, VSP\_001984;  
RX Name=PRL-R133;  
RN IsoId=Q08501-134; Sequence=VSP\_001985, VSP\_001986;  
RX Name=PRL-R134;  
RN IsoId=Q08501-135; Sequence=VSP\_001987, VSP\_001988;  
RX Name=PRL-R135;  
RN IsoId=Q08501-136; Sequence=VSP\_001989, VSP\_001990;  
RX Name=PRL-R136;  
RN IsoId=Q08501-137; Sequence=VSP\_001991, VSP\_001992;  
RX Name=PRL-R137;  
RN IsoId=Q08501-138; Sequence=VSP\_001993, VSP\_001994;  
RX Name=PRL-R138;  
RN IsoId=Q08501-139; Sequence=VSP\_001995, VSP\_001996;  
RX Name=PRL-R139;  
RN IsoId=Q08501-140; Sequence=VSP\_001997, VSP\_001998;  
RX Name=PRL-R140;  
RN IsoId=Q08501-141; Sequence=VSP\_001999, VSP\_002000;  
RX Name=PRL-R141;  
RN IsoId=Q08501-142; Sequence=VSP\_002001, VSP\_002002;  
RX Name=PRL-R142;  
RN IsoId=Q08501-143; Sequence=VSP\_002003, VSP\_002004;  
RX Name=PRL-R143;  
RN IsoId=Q08501-144; Sequence=VSP\_002005, VSP\_002006;  
RX Name=PRL-R144;  
RN IsoId=Q08501-145; Sequence=VSP\_002007, VSP\_002008;  
RX Name=PRL-R145;  
RN IsoId=Q08501-146; Sequence=VSP\_002009, VSP\_002010;  
RX Name=PRL-R146;  
RN IsoId=Q08501-147; Sequence=VSP\_002011, VSP\_002012;  
RX Name=PRL-R147;  
RN IsoId=Q08501-148; Sequence=VSP\_002013, VSP\_002014;  
RX Name=PRL-R148;  
RN IsoId=Q08501-149; Sequence=VSP\_002015, VSP\_002016;  
RX Name=PRL-R149;  
RN IsoId=Q08501-150; Sequence=VSP\_002017, VSP\_002018;  
RX Name=PRL-R150;  
RN IsoId=Q08501-151; Sequence=VSP\_002019, VSP\_002020;  
RX Name=PRL-R151;  
RN IsoId=Q08501-152; Sequence=VSP\_002021, VSP\_002022;  
RX Name=PRL-R152;  
RN IsoId=Q08501-153; Sequence=VSP\_002023, VSP\_002024;  
RX Name=PRL-R153;  
RN IsoId=Q08501-154; Sequence=VSP\_002025, VSP\_002026;  
RX Name=PRL-R154;  
RN IsoId=Q08501-155; Sequence=VSP\_002027, VSP\_002028;  
RX Name=PRL-R155;  
RN IsoId=Q08501-156; Sequence=VSP\_002029, VSP\_002030;  
RX Name=PRL-R156;  
RN IsoId=Q08501-157; Sequence=VSP\_002031, VSP\_002032;  
RX Name=PRL-R157;  
RN IsoId=Q08501-158; Sequence=VSP\_002033, VSP\_002034;  
RX Name=PRL-R158;  
RN IsoId=Q08501-159; Sequence=VSP\_002035, VSP\_002036;  
RX Name=PRL-R159;  
RN IsoId=Q08501-160; Sequence=VSP\_002037, VSP\_002038;  
RX Name=PRL-R160;  
RN IsoId=Q08501-161; Sequence=VSP\_002039, VSP\_002040;  
RX Name=PRL-R161;  
RN IsoId=Q08501-162; Sequence=VSP\_002041, VSP\_002042;  
RX Name=PRL-R162;  
RN IsoId=Q08501-163; Sequence=VSP\_002043, VSP\_002044;  
RX Name=PRL-R163;  
RN IsoId=Q08501-164; Sequence=VSP\_002045, VSP\_002046;  
RX Name=PRL-R164;  
RN IsoId=Q08501-165; Sequence=VSP\_002047, VSP\_002048;  
RX Name=PRL-R165;  
RN IsoId=Q08501-166; Sequence=VSP\_002049, VSP\_002050;  
RX Name=PRL-R166;  
RN IsoId=Q08501-167; Sequence=VSP\_002051, VSP\_002052;  
RX Name=PRL-R167;  
RN IsoId=Q08501-168; Sequence=VSP\_002053, VSP\_002054;  
RX Name=PRL-R168;  
RN IsoId=Q08501-169; Sequence=VSP\_002055, VSP\_002056;  
RX Name=PRL-R169;  
RN IsoId=Q08501-170; Sequence=VSP\_002057, VSP\_002058;  
RX Name=PRL-R170;  
RN IsoId=Q08501-171; Sequence=VSP\_002059, VSP\_002060;  
RX Name=PRL-R171;  
RN IsoId=Q08501-172; Sequence=VSP\_002061, VSP\_002062;  
RX Name=PRL-R172;  
RN IsoId=Q08501-173; Sequence=VSP\_002063, VSP\_002064;  
RX Name=PRL-R173;  
RN IsoId=Q08501-174; Sequence=VSP\_002065, VSP\_002066;  
RX Name=PRL-R174;  
RN IsoId=Q08501-175; Sequence=VSP\_002067, VSP\_002068;  
RX Name=PRL-R175;  
RN IsoId=Q08501-176; Sequence=VSP\_002069, VSP\_002070;  
RX Name=PRL-R176;  
RN IsoId=Q08501-177; Sequence=VSP\_002071, VSP\_002072;  
RX Name=PRL-R177;  
RN IsoId=Q08501-178; Sequence=VSP\_002073, VSP\_002074;  
RX Name=PRL-R178;  
RN IsoId=Q08501-179; Sequence=VSP\_002075, VSP\_002076;  
RX Name=PRL-R179;  
RN IsoId=Q08501-180; Sequence=VSP\_002077, VSP\_002078;  
RX Name=PRL-R180;  
RN IsoId=Q08501-181; Sequence=VSP\_002079, VSP\_002080;  
RX Name=PRL-R181;  
RN IsoId=Q08501-182; Sequence=VSP\_002081, VSP\_002082;  
RX Name=PRL-R182;  
RN IsoId=Q08501-183; Sequence=VSP\_002083, VSP\_002084;  
RX Name=PRL-R183;  
RN IsoId=Q08501-184; Sequence=VSP\_002085, VSP\_002086;  
RX Name=PRL-R184;  
RN IsoId=Q08501-185; Sequence=VSP\_002087, VSP\_002088;  
RX Name=PRL-R185;  
RN IsoId=Q08501-186; Sequence=VSP\_002089, VSP\_002090;  
RX Name=PRL-R186;  
RN IsoId=Q08501-187; Sequence=VSP\_002091, VSP\_002092;  
RX Name=PRL-R187;  
RN IsoId=Q08501-188; Sequence=VSP\_002093, VSP\_002094;  
RX Name=PRL-R188;  
RN IsoId=Q08501-189; Sequence=VSP\_002095, VSP\_002096;  
RX Name=PRL-R189;  
RN IsoId=Q08501-190; Sequence=VSP\_002097, VSP\_002098;  
RX Name=PRL-R190;  
RN IsoId=Q08501-191; Sequence=VSP\_002099, VSP\_002100;  
RX Name=PRL-R191;  
RN IsoId=Q08501-192; Sequence=VSP\_002101, VSP\_002102;  
RX Name=PRL-R192;  
RN IsoId=Q08501-193; Sequence=VSP\_002103, VSP\_002104;  
RX Name=PRL-R193;  
RN IsoId=Q08501-194; Sequence=VSP\_002105, VSP\_002106;  
RX Name=PRL-R194;  
RN IsoId=Q08501-195; Sequence=VSP\_002107, VSP\_002108;  
RX Name=PRL-R195;  
RN IsoId=Q08501-196; Sequence=VSP\_002109, VSP\_002110;  
RX Name=PRL-R196;  
RN IsoId=Q08501-197; Sequence=VSP\_002111, VSP\_002112;  
RX Name=PRL-R197;  
RN IsoId=Q08501-

DR SMART; SMO0060; FN3; 2;  
 DR PROSITE; PS01352; HEMATOPOREC L F1; 1;  
 KW Receptor; Transmembrane; Glycoprotein; Signal; Repeat;  
 KW Alternative splicing;  
 FT SIGNAL 1 19  
 FT CHAIN 20 608  
 FT DOMAIN 20 229  
 FT TRANSMEM 230 253  
 FT DOMAIN 254 608  
 FT DOMAIN 254 608  
 FT DOMAIN 119 222  
 FT DISULFID 31 41  
 FT DISULFID 70 81  
 FT CARBOHYD 54 54  
 FT CARBOHYD 99 99  
 FT CARBOHYD 127 127  
 FT VARSPLIC 281 292  
 FT FT  
 FT FT  
 FT VARSPLIC 293 608  
 FT VARSPLIC 281 303  
 FT VARSPLIC 304 608  
 FT CONFLICT 558 558  
 FT SEQUENCE 608 AA; 68240 MW; B8CE202B2EFC9FC6 CRC64;  
 Query Match 16.3%; Score 58.5; DB 1; Length 608;  
 Best Local Similarity 24.7%; Pred. No. 41;  
 Matches 18; Conservative 12; Mismatches 24; Indels 19; Gaps 2;  
 QY 2 SPTLTQSTNTHNTOSSSSSDGLFRSPAHSLLPGEDGRVEYVDFAEFRLWS----- 56  
 DB 388 TPNNQNTNCHTDSKST-----WPLPPGQHTRRSPYHSIADVCKLAGSPGDT 437  
 QY 57 ----VDHGEQSW 65  
 DB 438 LDSFLDKAEENVL 450  
 RESULT 14  
 AF17 HUMAN  
 ID AF17 HUMAN STANDARD; PRT; 1093 AA.  
 AC P55198;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE AF-17 protein.  
 GN MLLT6 OR AF17.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94336695; PubMed=8058765;  
 RA Prasad R., Leshkowitz D., Gu Y., Alder H., Nakamura T., Saito H.,  
 RA Ruebner K., Berger R., Croce C.M., Canaanani E.;  
 RT "Leucine-zipper dimerization motif encoded by the AF17 gene fused to  
 RT ALL-1 (MLL) in acute leukemia."  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:8107-8111(1994).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -!- DISEASE: INVOLVED IN ACUTE LEUKEMIAS BY A CHROMOSOMAL  
 CC TRANSLOCATION T(11;17)(Q23;Q21) THAT INVOLVES MLLT6 AND MLL/HEX.  
 CC THE RESULT IS A ROGUE ACTIVATOR PROTEIN.  
 CC -!- SIMILARITY: Contains 2 PHD-type zinc fingers.  
 CC -!- SIMILARITY: HIGH, TO AF10.  
 CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/AF17.html".  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U07932; AAA21145.1; -  
 DR PIR; I38533; I38533.  
 DR Genew; HGNC:7138; MLLT6.  
 DR MIM; 600328; -  
 DR GO; GO:0007048; P:oncogenesis; TAS.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; TAS.  
 DR InterPro; IPR001965; Znf\_PHD.  
 DR Pfam; PF00628; PHD; 1.  
 DR SMART; SMO0249; PHD; 2.  
 DR PROSITE; PS01359; ZF PHD 1; 1.  
 DR PROSITE; PS00016; ZF PHD 2; 2.  
 KW Nuclear protein; Zinc-finger; Repeat; Chromosomal translocation;  
 KW Proto-oncogene.  
 FT ZN FING 5 57 PHD-TYPE 1.  
 FT ZN FING 51 68 C4-TYPE.  
 FT ZN FING 117 180 PHD-TYPE 2.  
 FT DOMAIN 190 211 GLY/SER-RICH.  
 FT DOMAIN 275 282 POLY-SER.  
 FT DOMAIN 326 338 POLY-SER.  
 FT DOMAIN 729 764 LEUCINE-ZIPPER.  
 FT DOMAIN 822 829 POLY-SER.  
 FT DOMAIN 834 862 PRO-RICH.  
 FT DOMAIN 935 984 GLN-RICH.  
 FT DOMAIN 1040 1051 POLY-ALA.  
 FT DOMAIN 1069 1080 GLY-RICH.  
 FT SITE 551 551 MLL FUSION POINT (IN ACUTE MYELOID  
 FT LEUKEMIA PATIENT).  
 SQ SEQUENCE 1093 AA; 112021 MW; F60042A6D3BF579E CRC64;  
 Query Match 16.3%; Score 58.5; DB 1; Length 1093;  
 Best Local Similarity 36.4%; Pred. No. 84;  
 Matches 20; Conservative 8; Mismatches 18; Indels 9; Gaps 2;  
 QY 14 TQSSSSDGLFRSPAHSLLPGEDGRVEYVDFAEFRLWSVDHGEQSWVTAP 68  
 DB 330 SSSSSSSSGGPF--QPAVS-----SLQSSPDSAPPKLEQPEEDKYKPTAP 375  
 RESULT 15  
 PLC1 CANAL  
 ID PLC1 CANAL STANDARD; PRT; 1099 AA.  
 AC O13433;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase 1  
 DE (EC 3.1.4.11) (Phosphoinositide phospholipase C) (PLC-1)  
 DE (Phospholipase C-1).  
 GN PLC1.  
 OS Candida albicans (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
 OX NCBI\_TaxID=5476;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=132A;  
 RX MEDLINE=98129081; PubMed=9467900;  
 RA Bennett D.E., McCreary C.E., Coleman D.C.;  
 RT "Genetic characterization of a phospholipase C gene from Candida  
 RT albicans: presence of homologous sequences in Candida species other  
 RT than Candida albicans."  
 RL Microbiology 144:55-72(1998).  
 CC -!- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES  
 CC DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS  
 CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE  
 CC C ENZYMES.



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:33:11 ; Search time 21.1351 Seconds  
(without alignments)  
598.382 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSPILTQSTNTHQTSSSS.....AEFYRLMSVDHGQSVWTAP 68

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pbp:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pbp:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pbp:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pbp:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pbp:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pbp:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pbp:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pbp:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pbp:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pbp:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pbp:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pbp:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pbp:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pbp:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pbp:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pbp:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pbp:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pbp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	504	12	US-10-384-743-2
2	359	100.0	504	14	US-10-123-427-2
3	359	100.0	504	14	US-10-123-427-6
4	359	100.0	504	14	US-10-158-895-2
5	359	100.0	513	12	US-10-384-743-43
6	359	100.0	513	14	US-10-158-895-43
7	359	100.0	517	12	US-10-384-743-11
8	359	100.0	517	14	US-10-158-895-11
9	352	98.1	84	10	US-09-925-300-1270
10	77	21.4	16	12	US-10-384-743-41
11	77	21.4	16	14	US-10-158-895-41
12	69.5	19.4	199	10	US-09-941-831-21
13	67.5	18.8	739	12	US-10-374-979-89
14	67.5	18.8	739	15	US-10-097-534-10
15	67.5	18.8	752	11	US-09-919-039-235

16	66	18.4	16	12	US-10-384-743-40	Sequence 40, Appl
17	66	18.4	16	14	US-10-158-895-40	Sequence 40, Appl
18	65.5	18.2	498	14	US-10-037-667-5	Sequence 5, Appl
19	64.5	18.0	244	15	US-10-012-542-463	Sequence 463, App
20	63.5	17.7	240	9	US-09-853-161-100	Sequence 100, App
21	63.5	17.7	240	9	US-09-852-659A-100	Sequence 100, App
22	63.5	17.7	240	10	US-09-852-797-100	Sequence 100, App
23	63.5	17.7	335	9	US-09-853-161-64	Sequence 64, Appl
24	63.5	17.7	335	9	US-09-852-659A-64	Sequence 64, Appl
25	63.5	17.7	335	10	US-09-852-797-64	Sequence 64, Appl
26	63.5	17.7	824	12	US-10-210-951-58	Sequence 58, Appl
27	63.5	17.7	824	12	US-10-211-884-58	Sequence 58, Appl
28	63.5	17.7	824	15	US-10-226-844-1	Sequence 1, Appl
29	63.5	17.7	946	15	US-10-156-761-14486	Sequence 14486, A
30	63	17.5	850	9	US-09-915-181A-3	Sequence 3, Appl
31	62	17.3	141	12	US-10-029-386-30357	Sequence 30357, A
32	62	17.3	734	15	US-10-040-862-10463	Sequence 10463, A
33	61	17.0	322	9	US-09-764-853-606	Sequence 606, Appl
34	61	17.0	1119	15	US-10-245-802-12	Sequence 12, Appl
35	61	17.0	1511	10	US-09-801-368-250	Sequence 250, App
36	60	16.7	402	15	US-10-156-761-8505	Sequence 8505, Ap
37	60	16.7	1260	15	US-10-245-802-8	Sequence 8, Appl
38	58.5	16.3	235	9	US-09-799-777-43	Sequence 43, Appl
39	58.5	16.3	515	12	US-10-094-466-16	Sequence 16, Appl
40	58.5	16.3	3256	10	US-09-919-172-98	Sequence 98, Appl
41	58.5	16.3	3256	11	US-09-919-039-21	Sequence 21, Appl
42	58	16.2	692	15	US-10-156-761-12598	Sequence 12598, A
43	58	16.2	1047	15	US-10-245-802-22	Sequence 22, Appl
44	57.5	16.0	367	15	US-10-106-698-5815	Sequence 5815, Ap
45	57.5	16.0	394	9	US-09-815-242-10233	Sequence 10233, A

ALIGNMENTS

RESULT 1  
US-10-384-743-2  
; Sequence 2, Application US/10384743  
; Publication No. US20030162228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ IDS NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-384-743-2

Query Match 100.0%; Score 359; DB 12; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.6e-34;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 QSPILTQSTNTHQTSSSSSDGGLFRSRPAHSIPPGEDGRVVEYVDFAEYRLMSVDHG 60  
Db 437 QSPILTQSTNTHQTSSSSSDGGLFRSRPAHSIPPGEDGRVVEYVDFAEYRLMSVDHG 496  
Qy 61 BQSVWTAP 68  
Db 497 BQSVWTAP 504

## RESULT 2

US-10-123-427-2  
; Sequence 2, Application US/10123427  
; Publication No. US20020119525A1  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/123,427  
; FILING DATE: 17-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/406,854  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-123-427-2

Query Match 100.0%; Score 359; DB 14; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.6e-34;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEYRLWSVDHG 60  
Db 437 QSPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEYRLWSVDHG 496  
Qy 61 EQSVVTAP 68  
Db 497 EQSVVTAP 504

## RESULT 3

US-10-123-427-6  
; Sequence 6, Application US/10123427  
; Publication No. US20020119525A1  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR

NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/123,427  
FILING DATE: 17-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/406,854  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/08/752,891  
FILING DATE: 20-NOV-1996  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 504 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-123-427-6

Query Match 100.0%; Score 359; DB 14; Length 504;  
Best Local Similarity 100.0%; Pred. No. 1.6e-34;  
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEYRLWSVDHG 60  
Db 437 QSPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEYRLWSVDHG 496  
Qy 61 EQSVVTAP 68  
Db 497 EQSVVTAP 504

## RESULT 4

US-10-158-895-2  
; Sequence 2, Application US/10158895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188



```

; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-2

Query Match      100.0%; Score 359; DB 14; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.6e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPITLQSTNTHTCSSSSDGGFLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 60
Db 437 QSPITLQSTNTHTCSSSSDGGFLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 496
QY 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 5
US-10-384-743-43
; Sequence 43, Application US/10384743
; Publication No. US20030162228A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/384,743
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-384-743-43

Query Match      100.0%; Score 359; DB 12; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.6e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPITLQSTNTHTCSSSSDGGFLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 60
Db 446 QSPITLQSTNTHTCSSSSDGGFLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 505
QY 61 EQSVVTAP 68
Db 506 EQSVVTAP 513

RESULT 6
US-10-158-895-43
; Sequence 43, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03

```

```

; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-43

Query Match      100.0%; Score 359; DB 14; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.6e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPITLQSTNTHTCSSSSDGGFLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 60
Db 446 QSPITLQSTNTHTCSSSSDGGFLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 505
QY 61 EQSVVTAP 68
Db 506 EQSVVTAP 513

RESULT 7
US-10-384-743-11
; Sequence 11, Application US/10384743
; Publication No. US20030162228A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/384,743
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-384-743-11

Query Match      100.0%; Score 359; DB 12; Length 517;
Best Local Similarity 100.0%; Pred. No. 1.7e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPITLQSTNTHTCSSSSDGGFLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 60
Db 437 QSPITLQSTNTHTCSSSSDGGFLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 496
QY 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 8
US-10-158-895-11
; Sequence 11, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO

```

```
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-11

Query Match      100.0%; Score 359; DB 14; Length 517;
Best Local Similarity 100.0%; Pred. No. 1.7e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGDDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGDDGRVEPYVDFAEFYRLWSVDHG 496

Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 9
US-09-925-300-1270
; Sequence 1270, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1270
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1270

Query Match      98.1%; Score 352; DB 10; Length 84;
Best Local Similarity 98.5%; Pred. No. 1.3e-34;
Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QSPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGDDGRVEPYVDFAEFYRLWSVDHG 60
Db 17 QSPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGDDGRVEPYVDFAEFYRLWSVDHG 76

Qy 61 EQSVVTAP 68
Db 77 EQSVVTAP 84

RESULT 10
```

```
US-10-384-743-41
; Sequence 41, Application US/10384743
; Publication No. US2003016228A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/384,743
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-384-743-41

Query Match      21.4%; Score 77; DB 12; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPTLTQSTNTHQT 15
Db 2 QSPTLTQSTNTHQT 16

RESULT 11
US-10-158-895-41
; Sequence 41, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-158-895-41

Query Match      21.4%; Score 77; DB 14; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPTLTQSTNTHQT 15
Db 2 QSPTLTQSTNTHQT 16
```

## RESULT 12

US-09-941-831-21  
; Sequence 21, Application US/09941831  
; Patent No. US20020160493A1  
; GENERAL INFORMATION:  
; APPLICANT: Ebner et al.  
; TITLE OF INVENTION: PT049p1  
; CURRENT APPLICATION NUMBER: US/09/941,831  
; CURRENT FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: PCT/US01/06256  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/186,350  
; PRIOR FILING DATE: 2000-03-02  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 21  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-941-831-21

Query Match 19.4%; Score 69.5; DB 10; Length 199;  
Best Local Similarity 41.9%; Pred. No. 2.2;  
Matches 18; Conservative 3; Mismatches 17; Indels 5; Gaps 1;

QY 1 QSPTLTQSTNTHQTSSSSSDG-----LFRSRPAHSLPPGE 38  
Db 119 QSPTLTQSTNTHQTSSSSSDG-----LFRSRPAHSLPPGE 161

## RESULT 13

US-10-374-979-89  
; Sequence 89, Application US/10374979  
; Publication No. US20030219793A1  
; GENERAL INFORMATION:  
; APPLICANT: John P. Carulli et al.  
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3  
; FILE REFERENCE: 032796-021  
; CURRENT APPLICATION NUMBER: US/10/374,979  
; CURRENT FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: US 09/544,398  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 09/543,771  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 09/229,319  
; PRIOR FILING DATE: 1999-01-13  
; PRIOR APPLICATION NUMBER: US 60/071,449  
; PRIOR FILING DATE: 1998-01-13  
; PRIOR APPLICATION NUMBER: US 60/105,511  
; PRIOR FILING DATE: 1998-10-23  
; NUMBER OF SEQ ID NOS: 109  
; SEQ ID NO 89  
; LENGTH: 739  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-374-979-89

Query Match 18.8%; Score 67.5; DB 12; Length 739;  
Best Local Similarity 31.1%; Pred. No. 18;  
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

QY 2 SPTLTQSTNTHQTSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48  
Db 119 SSTGSLPPTNTNTSEGATSLIPLTISGSGRPLNPVTQAPLPFGWEQRVD----- 173  
QY 49 AEFYRLMSVDHGEQ 62  
Db 174 -QHGRVYVDHVEK 186

## RESULT 14

US-10-097-534-10  
; Sequence 10, Application US/10097534  
; Publication No. US20030049607A1  
; GENERAL INFORMATION:  
; APPLICANT: GREENER, TSVIKA  
; APPLICANT: MOSKOWITZ, HAIM  
; APPLICANT: REISS, YUVAL  
; APPLICANT: ALROY, IRIS  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL  
; TITLE OF INVENTION: MATURATION  
; FILE REFERENCE: PLV-001.01  
; CURRENT APPLICATION NUMBER: US/10/097,534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/275,224  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/308,958  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/340,170  
; PRIOR FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 739  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-534-10

Query Match 18.8%; Score 67.5; DB 15; Length 739;  
Best Local Similarity 31.1%; Pred. No. 18;  
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

QY 2 SPTLTQSTNTHQTSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48  
Db 119 SSTGSLPPTNTNTSEGATSLIPLTISGSGRPLNPVTQAPLPFGWEQRVD----- 173  
QY 49 AEFYRLMSVDHGEQ 62  
Db 174 -QHGRVYVDHVEK 186

## RESULT 15

US-09-919-039-235  
; Sequence 235, Application US/09919039  
; Publication No. US20030108871A1  
; GENERAL INFORMATION:  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES  
; FILE REFERENCE: PA-0035 US  
; CURRENT APPLICATION NUMBER: US/09/919,039  
; CURRENT FILING DATE: 2002-09-09  
; PRIOR APPLICATION NUMBER: 60/222,113  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 401  
; SOFTWARE: PERL Program  
; SEQ ID NO 235  
; LENGTH: 752  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1989186CD1  
US-09-919-039-235

Query Match 18.8%; Score 67.5; DB 11; Length 752;  
Best Local Similarity 31.1%; Pred. No. 19;  
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;  
QY 2 SPTLTQSTNTHQTSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48  
Db 132 SSTGSLPPTNTNTSEGATSLIPLTISGSGRPLNPVTQAPLPFGWEQRVD----- 186  
QY 49 AEFYRLMSVDHGEQ 62

Db : |:: ||| |:  
187 -QHGRVYVDHVEK 199

Search completed: December 4, 2003, 09:44:39  
Job time : 21.1351 secs



A;Residues: 1-579 <SAK>  
A;Cross-references: DDBJ:AB009356; NID:g2924623; PIDN:BA25025.1; PID:g2924624  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C;Keywords: phosphotransferase

Query Match 100.0%; Score 1252; DB 2; Length 579;  
Best Local Similarity 100.0%; Pred. No. 4.1e-73;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHPIVKLYGACLPVCLVMEYAEAGSLYNVLHGAELPYPTAAHAMSACL 60  
DB 76 VELRQLSRVNHPIVKLYGACLPVCLVMEYAEAGSLYNVLHGAELPYPTAAHAMSACL 135

QY 61 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHTMNNKGSAAW 120  
DB 136 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHTMNNKGSAAW 195

QY 121 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
DB 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTRCWSKDPQRSMEIIVKIMTHLMRYFFGADPELQYPCQ 228  
DB 256 PKPIESLMTRCWSKDPQRSMEIIVKIMTHLMRYFFGADPELQYPCQ 303

RESULT 3  
JC5956  
transforming growth factor-beta activated kinase (EC 2.7.1.-) 1b - human  
C;Species: Homo sapiens (man)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C;Accession: JC5956  
R;Sakurai, H.; Shigemori, N.; Hasegawa, K.; Saito, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
A;Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind  
A;Reference number: JC5955; MUID:98153801; PMID:9480845  
A;Accession: JC5956  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-606 <SAK>  
A;Cross-references: DDBJ:AB009357; NID:g2924625; PIDN:BA25026.1; PID:g2924626  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C;Keywords: phosphotransferase

Query Match 100.0%; Score 1252; DB 2; Length 606;  
Best Local Similarity 100.0%; Pred. No. 4.3e-73;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHPIVKLYGACLPVCLVMEYAEAGSLYNVLHGAELPYPTAAHAMSACL 60  
DB 76 VELRQLSRVNHPIVKLYGACLPVCLVMEYAEAGSLYNVLHGAELPYPTAAHAMSACL 135

QY 61 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHTMNNKGSAAW 120  
DB 136 QCSQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHTMNNKGSAAW 195

QY 121 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
DB 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTRCWSKDPQRSMEIIVKIMTHLMRYFFGADPELQYPCQ 228  
DB 256 PKPIESLMTRCWSKDPQRSMEIIVKIMTHLMRYFFGADPELQYPCQ 303

RESULT 4  
S68178  
mixed-lineage protein kinase 2 (EC 2.7.1.-) - human  
C;Species: Homo sapiens (man)  
C;Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 11-Jun-1999  
C;Accession: S68178; I38044; S32468  
R;Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simps  
Eur. J. Biochem. 234, 492-500, 1995

A;Title: Complete nucleotide sequence, expression, and chromosomal localisation of huma  
A;Reference number: S68178; MUID:96128179; PMID:8536694  
A;Accession: S68178  
A;Molecule type: mRNA  
A;Residues: 1-954 <DOR>  
A;Cross-references: EMBL:X90846; NID:g971419; PIDN:CAA62351.1; PID:g971420  
R;Kato, M.; Hirai, M.; Sugimura, T.; Terada, M.  
Oncogene 10, 1447-1451, 1995  
A;Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase  
A;Reference number: I38044; MUID:95249256; PMID:7731697  
A;Accession: I38044  
A;Status: translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-461, 'A', 'V', 465-470, 'S', 472-806, 'R', 808-817, 'A', 819-954 <RES>  
A;Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA8831.1; PID:g758593  
R;Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.  
Eur. J. Biochem. 213, 701-710, 1993  
A;Title: Identification of a new family of human epithelial protein kinases containing  
A;Reference number: S32467; MUID:93238756; PMID:8477742  
A;Accession: S32468  
A;Molecule type: mRNA  
A;Residues: 244-464, 'AQAAGRRQHPQALWL' <DO2>  
C;Genetics:  
A;Gene: GDB:MLK2; GDB:MST  
A;Cross-references: GDB:362654; GDB:624810; OMIM:600137  
A;Map position: 19q13.1-19q13.2  
C;Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology  
C;Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein  
F;23-76/Domain: SH3 homology <SH3>  
F;96-364/Domain: protein kinase homology <KIN>  
F;104-112/Region: protein kinase ATP-binding motif  
F;384-405/Region: leucine zipper motif  
F;419-440/Region: leucine zipper motif  
F;449-463/Region: basic  
F;125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 31.3%; Score 392.5; DB 1; Length 954;  
Best Local Similarity 40.6%; Pred. No. 5.4e-18;  
Matches 88; Conservative 35; Mismatches 77; Indels 17; Gaps 6;

QY 2 ELRQLSRVNHPIVKLYGACLP--VCLVMEYAEAGSLYNVLHGAELPYPTAAHAMSACL 59  
DB 145 EARLFGALQHPNIALRGACLP--VCLVMEYAEAGSLYNVLHGAELPYPTAAHAMSACL 200

QY 60 LQCSQGVAYLHSMOPKALIHRLDKPPNLLV-----AGTVLKICDFGTACDIQTHTM 111  
DB 201 VQVARGMNYLHNDAPVPIIHRDLKSNILILEAIEHNENLADTVLKITDFGLAREWHKTK 260

QY 112 TNKKGSAWNAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAV-HN 170  
DB 261 MSAGTYAWNAPEVIRLSFKSDVNSFGVLWELLTGVEPYREL--DALAVAYGVAMN 318

QY 171 GTRPPLIKNLKPPIESLMTRCWSKDPQRSMEIIVK 207  
DB 319 KLTLPIPTCTCEFFARLLEECWDPDPHGRDPFGSILK 355

RESULT 5  
G84635  
probable protein kinase [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002  
C;Accession: G84635  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: G84635  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-407 <STO>

A;Cross-references: GB:AE002093; NID:G4337195; PIDN:AAD18109.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: At2G24360  
A;Map position: 2  
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 29.9%; Score 374; DB 2; Length 407;  
Best Local Similarity 39.4%; Pred. No. 3.9e-17;  
Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHPIVLYGACLNVP--CLVMEYAEAGSGLYNVL-----HGAELPYPYTAHA 55  
Db EVSMLANLKHPIVFIAGACRPMWCIIVTEYAKGVSQFLTRQNRAVPLKL-----A 229  
QY 56 MSWCLQCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFTA-CDIQTH-MTN 113  
Db VQKALDVARGWAYVHG---RNFIHRLKSDNLLISADKSI-KIADFGVARIIEVQTEGMP 285  
QY 114 NKGSAAMWAPVFEAGSYSEKCDVFSWGIILWEVITRRKPFDEIGG--PAPRIMWAVHG 171  
Db ETGTYRWMAPEMIQHRAYNQKVDVYSGVILWELITGLLPFQNTAVQAFAV---VNRG 342  
QY 172 TRPPLIKNLKPISLMTKRCWKDPSQSPSMEEIVKIM 209  
Db VRPTVPNDCLPVLSDIMTRCWDANPEVRPCFVEVVKLL 380

## RESULT 6

T10671  
protein kinase homolog F6E21.90 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 19-May-2000  
C;Accession: T10671  
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z16533  
A;Accession: T10671  
A;Molecule type: DNA  
A;Residues: 1-412 <BEV>  
A;Cross-references: ENBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90  
A;Experimental source: cultivar Columbia; BAC clone F6E21  
C;Genetics:  
A;Gene: ATSP:F6E21.90  
A;Map position: 4  
A;Introns: 300/2  
C;Superfamily: kinase-related transforming protein; protein kinase homology  
F;135-392/Domain: protein kinase homology <KIN>

Query Match 29.8%; Score 373; DB 2; Length 412;  
Best Local Similarity 39.9%; Pred. No. 4.6e-17;  
Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHPIVLYGACLNVP--CLVMEYAEAGSGLYNVL-----HGAELPYPYTAHA 55  
Db 180 EVSMLANLKHPIVFIAGACRPMWCIIVTEYAKGVSQFLTRQNRAVPLKL-----A 234  
QY 56 MSWCLQCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFTA-CDIQTH-MTN 113  
Db 235 VQKALDVARGWAYVHE---RNFIHRLKSDNLLISADKSI-KIADFGVARIIEVQTEGMP 290  
QY 114 NKGSAAMWAPVFEAGSYSEKCDVFSWGIILWEVITRRKPFDEIGG--PAPRIMWAVHG 171  
Db 291 ETGTYRWMAPEMIQHRPYTKVDVYSGVILWELITGLLPFQNTAVQAFAV---VNRG 347  
QY 172 TRPPLIKNLKPISLMTKRCWKDPSQSPSMEEIVKIM 209  
Db 348 VRPTVPNDCLPVLGEIMTRCWDADPEVRPCFAEIVNLL 385

## RESULT 7

JU0229  
mixed-lineage protein kinase 1 - human  
C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999  
C;Accession: S32467; JU0229  
R;Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.  
Eur. J. Biochem. 213, 701-710, 1993  
A;Title: Identification of a new family of human epithelial protein kinases containing t  
A;Reference number: S32467; MUID:93238756; PMID:8477742  
A;Accession: S32467  
A;Molecule type: mRNA  
A;Residues: 1-394 <DO2>

C;Genetics:  
A;Gene: GDB:MLK1  
A;Cross-references: GDB:141921; OMIM:600136  
A;Map position: 14q24.3-14q31  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C;Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k  
F;1-269/Domain: protein kinase homology <KIN>  
F;1-268/Domain: catalytic <CAT>  
F;9-17/Region: protein kinase ATP-binding motif  
F;289-310/Region: leucine zipper motif  
F;324-345/Region: leucine zipper motif  
F;354-368/Region: basic

Query Match 29.0%; Score 363.5; DB 2; Length 394;  
Best Local Similarity 37.7%; Pred. No. 1.8e-16;  
Matches 83; Conservative 36; Mismatches 84; Indels 17; Gaps 6;

QY 2 ELRQLSRVNHPIVLYGACLNVP--CLVMEYAEAGSGLYNVLHGAELPYPYTAHAHMSWC 59  
Db 50 EAKLFAMLKHPNITIALRGVCLKEPNLCVMEFARGGPLNRLVLSGKRIPPD1----LVNWA 105  
QY 60 LQCSQGVAYLHSMQPKALIHRLKPPNLLV-----AGTVLKI CDFTACD-IQTHM 111  
Db 106 VQIARGMNYLHDEAIVPIIHRDLKSSNLIILQKVENGLSNKILKITDFGLAREWHRTTK 165  
QY 112 TNNKSAAMWAPVFEAGSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HN 170  
Db 166 MSAAGTYAWMAPEVIRASMFSGSDVMSYGVLLNELLTGEVFFRGIDG--LRVAYGVAMN 223  
QY 171 GTRPPLIKNLKPISLMTKRCWKDPSQSPSMEEIVKIMT 210  
Db 224 KLALPISTCTPEPAFLKMEDCNPDPSRPSFTNILDQLT 263

## RESULT 8

T18287  
protein-tyrosine kinase (EC 2.7.1.112) - slime mold (Dictyostelium discoideum)  
C;Species: Dictyostelium discoideum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Feb-2000  
C;Accession: T18287  
R;Adler, K.  
submitted to the EMBL Data Library, July 1996  
A;Reference number: Z18856  
A;Accession: T18287  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1338 <ADL>  
A;Cross-references: EMBL:U64830; NID:G1468983; PID:G1468983; PIDN:AAE04999.1  
C;Genetics:  
A;Introns: 1181/3  
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match 28.9%; Score 362; DB 2; Length 1338;  
Best Local Similarity 38.1%; Pred. No. 6.4e-16;  
Matches 80; Conservative 41; Mismatches 73; Indels 16; Gaps 6;

QY 2 ELRQLSRVNHPIVLYGACLNVPCLVMEYAEAGSGLYNVLH-----GAELPYPYTAHMS 57  
Db 1100 EVSSLKSHENPVVTFMGARIDPPCIFTEYIQGSLYDLVLIHQIKLNLPLMWYKWIHDL- 1158  
QY 58 WCLQCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFTACDIQTHMT-NNKG 116  
Db 1159 -----SLGMEHLHSIQ---MLHRDLTSKNILLDEFFKNI-KIADFGLATTLSDMTLSGIT 1209





QY 177 IKNLKPPIESLMTRCWSKDPQSPQSMEEIVKIMTHL 212  
Db 332 PSSCPDGFKILLRQCWNSKPRNRPFRQ---ILLHL 364

RESULT 12  
A55318  
serine/threonine protein kinase (EC 2.7.1.1) DLK - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 24-Sep-1999  
C:Accession: A55318  
R:Holzman, L.B.; Merritt, S.E.; Fan, G.  
J. Biol. Chem. 269, 30808-30817, 1994  
A:Title: Identification, molecular cloning, and characterization of dual leucine zipper  
s  
A:Reference number: A55318; MUID:95074107; PMID:7983011  
A:Accession: A55318  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-888 <HOL>  
A:Cross-references: GB:U14636; NID:g602677; PIDN:AA57280.1; PID:g602678  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; leucine zipper; phosphotransferase  
F:156-404/Domain: protein kinase homology <KIN>  
F:164-172/Region: protein kinase ATP-binding motif

Query Match 28.2%; Score 353.5; DB 2; Length 888;  
Best Local Similarity 34.7%; Pred. No. 1.6e-15;  
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;

QY 2 ELRQLSRVNHNPVIVKLYGACLNPP--VCLVMEYAEAGSLYNVLHGAEPPLPYTAHAMSWC 59  
Db 194 DIKHLRLKLPNIITFGVGTQAPCYCILMEFCAQGQYEVLRAGRPV--TPSLLDVDS 250

QY 60 LQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK--GS 117  
Db 251 MGIAAGMNYLHLK---IIHRDLKSPN-MLITYDDVVKISDFGTSKLSKSTKNSFAGT 306

QY 118 AAWMAPEVIRNEPVSEKVDIWSFGVLLWELLTGTETPKYKDVDSA--IIWGVGSNSLHLPV 364  
Db 307 VAWMAPEVIRNEPVSEKVDIWSFGVLLWELLTGTETPKYKDVDSA--IIWGVGSNSLHLPV 364

QY 177 IKNLKPPIESLMTRCWSKDPQSPQSMEEIVKIMTHL 212  
Db 365 PSSCPDGFKILLRQCWNSKPRNRPFRQ---ILLHL 397

RESULT 13  
JC5399  
dual leucine zipper kinase (EC 2.7.-.-) - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 07-Jul-1997 #sequence\_revision 18-Jul-1997 #text\_change 28-May-1999  
C:Accession: JC5399  
R:Matsu, N.; Sarkar, G.; Shuto, T.; Marrs, J.; Bronk, J.T.; Mizuno, K.; Bolander, M.E.  
Biochem. Biophys. Res. Commun. 229, 571-576, 1996  
A:Title: Identification of a dual leucine zipper kinase involved in rat fracture repair.  
A:Reference number: JC5399; MUID:97127443; PMID:8954939  
A:Accession: JC5399  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-888 <MAT>  
C:Comment: This enzyme is involved in regulating cell function in the musculoskeletal system  
C:Genetics:  
A:Gene: rdlk  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; phosphotransferase  
F:54-75,89-98/Region: glycine-rich  
F:156-454/Domain: kinase catalytic #status predicted <CAT>  
F:164-404/Domain: protein kinase homology <KIN>  
F:164-172/Region: protein kinase ATP-binding motif  
F:421-449/Region: leucine zipper motif  
F:472-500/Region: leucine zipper motif

F:557-888/Region: glycine-serine-proline rich #status predicted

Query Match 28.1%; Score 351.5; DB 2; Length 888;  
Best Local Similarity 34.3%; Pred. No. 2.1e-15;  
Matches 74; Conservative 44; Mismatches 81; Indels 17; Gaps 8;

QY 2 ELRQLSRVNHNPVIVKLYGACLNPP--VCLVMEYAEAGSLYNVLHGAEPPLPYTAHAMSWC 59

Db 194 DIKHLRLKLPNIITFGVGTQAPCYCILMEFCAQGQYEVLRAGRPV--TPSLLDVDS 250

QY 60 LQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK--GS 117

Db 251 MGIAAGMNYLHLK---IIHRDLKSPN-MLITYDDVVKISDFGTSKLSKSTKNSFAGT 306

QY 118 AAWMAPEVIRNEPVSEKVDIWSFGVLLWELLTGTETPKYKDVDSA--IIWGVGSNSLHLPV 176

Db 307 VAWMAPEVIRNEPVSEKVDIWSFGVLLWELLTGTETPKYKDVDSA--IIWGVGSNSLHLPV 364

QY 177 IKNLKPPIESLMTRCWSKDPQSPQSMEEIVKIMTHL 212

Db 365 PSSCPDGFKILLRQCWNSKPRNRPFRQ---ILLHL 397

RESULT 14

D84555

probable protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C:Accession: D84555

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84555

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-546 <STO>

A:Cross-references: GB:AR02093; NID:96598802; PIDN:AAB80785.2; GSPDB:GNO0139

A:Gene: At2g17700

A:Map position: 2

Query Match 28.0%; Score 350; DB 2; Length 546;  
Best Local Similarity 36.3%; Pred. No. 1.7e-15;  
Matches 81; Conservative 41; Mismatches 83; Indels 18; Gaps 8;

QY 3 LRQLS-----RVNHPNIVKLYGACLN--PVCLVMEYAEAGSLYNVLHGAEPPLPYTAA 53

Db 325 LREFSQEVFIMRKVRHKNVQFLGACTRSPTLCIVTFEMARGSIYDFLHKQKCA--FKLQ 382

QY 54 HAMSCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDGTGTA-CDIQTH-M 111

Db 383 TLLKVALVDVAKGMYLHLQ---NNIIHRDLKTFANLMDHEGLV-KVADFGVARVQIESGM 438

QY 112 TNNKGSAAWMAPEVIRNEPVSEKVDIWSFGVLLWELLTGTETPKYKDVDSA--IIWGVGSNSLHLPV 171

Db 439 TAETGTVRMAPEVIRNEPVSEKVDIWSFGVLLWELLTGTETPKYKDVDSA--IIWGVGSNSLHLPV 497

QY 172 TRPLIKNLKPPIESLMTRCWSKDPQSPQSMEEIVKIMTHL 214

Db 498 LRPKPKPKTHPKVKGLLERCWHQDPEQRPLFEIEMLQIQIMK 540

RESULT 15

T48115

protein kinase ATMRL1 (EC 2.7.1.1) [imported] - Arabidopsis thaliana

N;Alternate names: protein F16M2.110

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2001

C:Accession: T48115; T51942

R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.

us-09-830-144-2 copy 76 303.rpr

search completed: Dec 20 1994  
Job time : 40.2838 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:08:45 ; Search time 77.7973 Seconds  
(without alignments)  
465.179 Million cell updates/sec

Title: US-09-830-144-2 COPY 76 303

perfect score: 1252

Sequence: 1 VELRQLSRVNHHPNIVKLYGA.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62

Gap0 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Minimum DB seq	length: 0
Maximum DB seq	length: 2000000000

post-processing: Minimum Match 0%

Post-processing: Minimum Match 0%  
Maximum Match 100%

Database :

A Genesee 19Jun03: \*

1:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1980.DAT.*
2:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1981.DAT.*
3:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1982.DAT.*
4:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1983.DAT.*
5:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1984.DAT.*
6:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1985.DAT.*
7:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1986.DAT.*
8:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1987.DAT.*
9:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1988.DAT.*
10:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1989.DAT.*
11:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1990.DAT.*
12:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1991.DAT.*
13:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1992.DAT.*
14:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1993.DAT.*
15:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1994.DAT.*
16:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1995.DAT.*
17:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1996.DAT.*
18:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1997.DAT.*
19:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1998.DAT.*
20:	/SIDSI/gcdata/genseq/genseqp-emb1/AA1999.DAT.*
21:	/SIDSI/gcdata/genseq/genseqp-emb1/AA2000.DAT.*
22:	/SIDSI/gcdata/genseq/genseqp-emb1/AA2001.DAT.*
23:	/SIDSI/gcdata/genseq/genseqp-emb1/AA2002.DAT.*
24:	/SIDSI/gcdata/genseq/genseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARY

Result No.	Score	Query Match	Length	DB	ID	Description
1	1252	100.0	567	20	AAV28998	Human TGF-beta act
2	1252	100.0	579	18	AAW27092	Mouse transforming
3	1252	100.0	579	20	AAW27093	Human transforming
4	1252	100.0	579	20	AAV28996	Human TGF-beta act
5	1252	100.0	579	20	AAV09542	Human TAK1 protein
6	1252	100.0	579	21	AAV91000	Human TAK-1 protei
7	1252	100.0	579	23	ABB5033	Pain regulated pro
8	1252	100.0	590	20	AAV09547	Human TAK1-6xHis p
9	1252	100.0	606	20	AAV28997	Human TGF-beta act

## ALIGNMENTS

## RESULT 1

AAV28998

ID AAY28998 standard; Protein; 567 AA.

XX

AC AAY28998;

XXX

DT. 29-OCT-1999 (first entry)

XX

DE Human TGF-beta activated kinase (TAK) 1c amino acid sequence.

XX

KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;

TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation; KW

KW intractable disease; atrophic dermatitis; psoriasis; viral infection;  
KW endotoxin shock; corticemia; human: btxvlg

endotoxin sho

XX  
Homocarpione

US Homo sapiens.  
yy

XX  
PN  
W09940202-A1

PN 009940202-AT.  
XX

12-AUG-1999.

XX  
FD-302 (Rev. 1-25-60)

02-FEB-1999:

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

XX Nuclear factor kappa B activation inhibitors, useful as preventives  
PT for, e.g. autoimmune diseases  
XX  
XX Examples; Page 43-46; 49pp; Japanese.  
XX  
XX The invention provides a method for identifying or screening a nuclear  
CC factor kappa B (NF-kB) activating inhibitor by examining the effect of a  
CC test substance on modulating the function(s) of TGF-beta activated kinase  
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
CC intractable diseases with inflammation (such as atrophic dermatitis and  
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The  
CC present sequence represents the amino acid sequence of human TAK1C  
CC (hTAK1C) protein.

XX SQ Sequence 567 AA;  
Query Match 100.0%; Score 1252; DB 20; Length 567;  
Best Local Similarity 100.0%; Pred. No. 3e-127;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VELRQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHMSWCL 60  
DB 76 VELRQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHMSWCL 135  
QY 61 QCSQGVAYLHSMOPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHTMTNKGSAW 120  
DB 136 QCSQGVAYLHSMOPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHTMTNKGSAW 195  
QY 121 MAPEVEGSGNSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
DB 196 MAPEVEGSGNSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255  
QY 181 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPFGADEPLQYPCQ 228  
DB 256 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPFGADEPLQYPCQ 303

RESULT 2  
AAW27092  
ID AAW27092 standard; Protein; 579 AA.  
AC AAW27092;  
XX  
XX 19-NOV-1997 (first entry)  
XX  
XX Mouse transforming growth factor-beta activated kinase TAK-1.  
XX  
XX TGF-beta; signal transmission; TGF-beta activated kinase;  
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;  
KW protein kinase.  
XX  
XX Mus musculus.  
XX  
XX JP09163990-A.  
XX 24-JUN-1997.  
XX  
XX 27-SEP-1996; 96JP-0256747.  
XX  
XX 24-JUL-1996; 96US-0685625.  
PR 29-SEP-1995; 95JP-0253549.  
XX  
XX (CHUS ) CHUGAI PHARM CO LTD.  
PA (UENO/) UENO N.  
XX  
XX WPI; 1997-380171/35.  
DR N-PSDB; AAT85094.  
XX  
XX DNA encoding transforming growth factor-beta-activated kinase, TAK-1  
PT - useful for studying the TGF-beta signal transmission system  
XX

PS Claim 14; Page 10-12; 20pp; Japanese.  
XX  
XX The present sequence represents mouse transforming growth factor-beta  
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the  
CC TAK-1 protein which is involved in the TGF-beta family signal  
CC transmission system. TAK-1, also known as activator of MAPK Kinase  
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone  
CC morphogenetic protein (BMP) and activates MAPK kinase by  
XX phosphorylation.  
XX SQ Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 18; Length 579;  
Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VELRQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHMSWCL 60  
DB 76 VELRQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHMSWCL 135  
QY 61 QCSQGVAYLHSMOPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHTMTNKGSAW 120  
DB 136 QCSQGVAYLHSMOPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHTMTNKGSAW 195  
QY 121 MAPEVEGSGNSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
DB 196 MAPEVEGSGNSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255  
QY 181 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPFGADEPLQYPCQ 228  
DB 256 PKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPFGADEPLQYPCQ 303

RESULT 3  
AAW27093  
ID AAW27093 standard; Protein; 579 AA.  
XX  
XX AAW27093;  
XX  
XX 19-NOV-1997 (first entry)  
XX  
XX Human transforming growth factor-beta activated kinase TAK-1.  
XX  
XX TGF-beta; signal transmission; TGF-beta activated kinase;  
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;  
KW protein kinase.  
XX  
XX Homo sapiens.  
XX  
XX JP09163990-A.  
XX 24-JUN-1997.  
XX  
XX 27-SEP-1996; 96JP-0256747.  
XX  
XX 24-JUL-1996; 96US-0685625.  
PR 29-SEP-1995; 95JP-0253549.  
XX  
XX (CHUS ) CHUGAI PHARM CO LTD.  
PA (UENO/) UENO N.  
XX  
XX WPI; 1997-380171/35.  
DR N-PSDB; AAT85095.  
XX  
XX DNA encoding transforming growth factor-beta-activated kinase, TAK-1  
PT - useful for studying the TGF-beta signal transmission system  
XX  
XX Claim 15; Page 13-15; 20pp; Japanese.  
XX  
XX The present sequence represents human transforming growth factor-beta  
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the  
CC TAK-1 protein which is involved in the TGF-beta family signal  
CC transmission system. TAK-1, also known as activator of MAPK Kinase  
CC

CC (AMK-1), is an enzyme which is activated by TGF-beta and bone  
CC morphogenetic protein (BMP) and activates MAPK kinase by  
CC phosphorylation.

XX  
SQ Sequence 579 AA;  
Query Match 100.0%; Score 1252; DB 18; Length 579;  
Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VELQLSRVNHPIVKLYGACINPVCLVMEYAEGSLYNVLHGAEPLPYTTAAHAMSACL 60  
DB 76 VELQLSRVNHPIVKLYGACINPVCLVMEYAEGSLYNVLHGAEPLPYTTAAHAMSACL 135  
QY 61 QCSQGVAVLHSMQPKALHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120  
DB 136 QCSQGVAVLHSMQPKALHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195  
QY 121 MAPEVFGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
DB 196 MAPEVFGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255  
QY 181 PKPIESLMTRCWSKDPSPQSPSMEIEIVKIMTHLMRYFPGADEPLQYPCQ 228  
DB 256 PKPIESLMTRCWSKDPSPQSPSMEIEIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 4  
AAY28996  
ID AAY28996 standard; Protein; 579 AA.

XX  
AC AAY28996;  
XX  
DT 29-OCT-1999 (first entry)  
XX  
DE Human TGF-beta activated kinase (TAK) 1a amino acid sequence.  
XX  
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;  
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;  
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;  
KW endotoxin shock; septicemia; human; hTAK1a.

XX  
OS Homo sapiens.  
XX  
PN WO9940202-A1.  
XX  
PD 12-AUG-1999.  
XX  
PF 02-FEB-1999; 99WO-JP00422.  
XX  
PR 30-OCT-1998; 98JP-0309316.  
PR 06-FEB-1998; 98JP-0026003.

XX  
FA (TANA ) TANABE SEIYAKU CO.  
XX  
PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;  
XX  
XX WPI; 1999-494298/41.  
DR N-ESDB; AAX99696.

XX  
PT Nuclear factor kappa B activation inhibitors, useful as preventives  
PT for, e.g. autoimmune diseases

XX  
PS Examples; Page 35-39; 49pp; Japanese.  
XX  
CC The invention provides a method for identifying or screening a nuclear  
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a  
CC test substance on modulating the function(s) of TGF-beta activated kinase  
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
CC intractable diseases with inflammation (such as atrophic dermatitis and  
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The  
CC present sequence represents the amino acid sequence of human TAK1a

CC (hTAK1a) protein.  
XX  
SQ Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 20; Length 579;  
Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VELQLSRVNHPIVKLYGACINPVCLVMEYAEGSLYNVLHGAEPLPYTTAAHAMSACL 60  
DB 76 VELQLSRVNHPIVKLYGACINPVCLVMEYAEGSLYNVLHGAEPLPYTTAAHAMSACL 135  
QY 61 QCSQGVAVLHSMQPKALHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120  
DB 136 QCSQGVAVLHSMQPKALHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195  
QY 121 MAPEVFGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
DB 196 MAPEVFGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255  
QY 181 PKPIESLMTRCWSKDPSPQSPSMEIEIVKIMTHLMRYFPGADEPLQYPCQ 228  
DB 256 PKPIESLMTRCWSKDPSPQSPSMEIEIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 5  
AAY09542  
ID AAY09542 standard; Protein; 579 AA.  
XX  
AC AAY09542;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE Human TAK1 protein.  
XX  
KW Human; TAK1; TAK1; screening; inhibition; TGF-beta;  
KW transforming growth factor beta.

XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 183..1922  
FT /\*tag= a  
XX  
PN WO9921010-A1.

XX  
PD 29-APR-1999.  
XX  
PF 22-OCT-1998; 98WO-JP04796.  
XX  
PR 22-OCT-1997; 97JP-0290188.  
XX  
PA (CHUS ) CHUGAI SEIYAKU KK.  
XX  
PI Ohtomo T, Ono K, Tsuchiya M;  
XX  
XX WPI; 1999-312645/26.  
DR N-ESDB; AAX56279.

XX  
PT Screening for TGF-beta inhibitory substances, which are useful as  
PT drugs for treatment of diseases relating to its disorder

XX  
PS Claim 4; Page 155-157; 195pp; Japanese.

XX  
CC A method has been developed for screening for substances which inhibit  
CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method  
CC comprises: (a) contacting the polypeptide in the presence of a sample;  
CC and (b) detecting the amount of bound polypeptide, in which the sample  
CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming  
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
CC indications e.g. as TGF-beta signal transduction inhibitors or  
CC activators, or extracellular matrix protein production enhancement  
CC inhibitors or activators, or cell proliferation prevention inhibitors or

CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence represents human TAK1.  
 XX  
 SQ Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 20; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHPIVLYGACLPVCLVMEYAEAGSLYNVLHGAELPYPTAAHMSWCL 60  
 DB 76 VELRQLSRVNHPIVLYGACLPVCLVMEYAEAGSLYNVLHGAELPYPTAAHMSWCL 135  
 QY 61 QCSQGVAYLHSMQPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHTMNNKGSAAW 120  
 DB 136 QCSQGVAYLHSMQPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHTMNNKGSAAW 195  
 QY 121 MAPEVFEGSNSYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
 DB 196 MAPEVFEGSNSYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL 255

QY 181 PKPIESLMTRCWSKDPQSPQSMSEIIVKIMTHLMRYFPFGADEPLQYPCQ 228  
 DB 256 PKPIESLMTRCWSKDPQSPQSMSEIIVKIMTHLMRYFPFGADEPLQYPCQ 303

RESULT 6  
 AAY91000  
 ID AAY91000 standard; Protein; 579 AA.

XX AC AAY91000;  
 XX AC AAY91000 (first entry)  
 DT 04-SEP-2000  
 XX Human TAK-1 protein sequence SEQ ID NO:2.  
 XX Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;  
 KW screening; signal transduction; inhibition; inflammatory cytokine;  
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;  
 KW antiinflammatory; suppression.  
 XX Homo sapiens.  
 OS WO200023610-A1.  
 XX 27-APR-2000.  
 XX 21-OCT-1999; 99WO-JP05817.  
 XX 21-OCT-1998; 98JP-0239962.  
 XX (CHUS) CHUGAI SEIYAKU KK.  
 XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;  
 XX WPI; 2000-339707/29.  
 XX N-PSDB; AAA39105.  
 XX Method for screening inhibitors of TAK1 signal transduction for  
 PT suppression of inflammatory cytokine production and use as  
 PT antiinflammatory agents -  
 XX Example 1; Page 80-84; 100pp; Japanese.

XX The present invention describes a method for screening compounds for  
 CC inhibition of inflammatory cytokine signal transduction by contacting  
 CC the sample with TAK1 and its receptor TAK1 and selecting for inhibition  
 CC of TAK1/TAK1 binding. Also described is a method for screening compounds  
 CC for inhibition of inflammatory cytokine signal transduction in which the

CC inhibition of TAK1 phosphorylation is selected for; and drug  
 CC compositions for the treatment of inflammatory disorders containing as  
 CC active component an inflammatory cytokine signal transduction inhibitor.  
 CC TAK1 is an essential component of the signalling process which results  
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),  
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used  
 CC for the selection of effective antiinflammatory agents. The present  
 CC sequence represents human TAK-1, which is used in the exemplification of  
 CC the present invention.  
 XX

SQ Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 21; Length 579;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHPIVLYGACLPVCLVMEYAEAGSLYNVLHGAELPYPTAAHMSWCL 60  
 DB 76 VELRQLSRVNHPIVLYGACLPVCLVMEYAEAGSLYNVLHGAELPYPTAAHMSWCL 135  
 QY 61 QCSQGVAYLHSMQPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHTMNNKGSAAW 120  
 DB 136 QCSQGVAYLHSMQPKALIHRLDLPNLLVAGGTVLKICDFGTACDIQTHTMNNKGSAAW 195  
 QY 121 MAPEVFEGSNSYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
 DB 196 MAPEVFEGSNSYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL 255  
 QY 181 PKPIESLMTRCWSKDPQSPQSMSEIIVKIMTHLMRYFPFGADEPLQYPCQ 228  
 DB 256 PKPIESLMTRCWSKDPQSPQSMSEIIVKIMTHLMRYFPFGADEPLQYPCQ 303

RESULT 7  
 ABB85033  
 ID ABB85033 standard; Protein; 579 AA.

XX AC ABB85033;  
 XX AC ABB85033 (first entry)  
 DT 16-MAY-2002  
 XX Pain regulated protein sequence 28.  
 DE Pain; analgesic; gene therapy; neurological disorder;  
 KW neurodegenerative disease.  
 XX Homo sapiens.  
 OS WO200212338-A2.  
 XX 14-FEB-2002.  
 XX 03-AUG-2001; 2001WO-EP09011.  
 XX 03-AUG-2000; 2000DE-1037759.  
 XX (CHEF) GRUENENTHAL GMBH.  
 XX Gillen C, Wetzel I, Whendt S, Weihe E, Schaefer MK;  
 XX WPI; 2002-257469/30.  
 XX N-PSDB; ABL88437.  
 XX Identifying pain-regulating compounds, useful for treating chronic pain  
 PT and for diagnosis, by measuring binding of compounds to specific  
 PT peptides and proteins -  
 XX Claim 1; Fig 44; 213pp; German.

XX The invention relates to identifying pain-regulating substances (A)  
 CC comprises (i) incubating a test substance with a cell (or preparation  
 CC from it) that has synthesised a peptide or protein (B) and (ii) measuring  
 CC either binding of the test substance to (B) or some functional parameter

that is altered by this binding. The method is useful for identifying pain-regulating substances (A) with analgesic activity, (A) along with nucleic acid (AB188411-AB188441) that encode proteins (B), AB95006-AB95037) that interact with (A); (B); vectors containing the nucleic acid; antibodies against (B); cells that express (B) and agents that bind to (B), are all useful for treating pain, particularly chronic pain, including use in gene therapy. The same materials can also be used for diagnosis, e.g. of neurological and neurodegenerative diseases. The present sequence is that of a polypeptide of the invention.

Sequence 579 AA;  
Query Match 100.0%; Score 1252; DB 23; Length 579;  
Best Local Similarity 100.0%; Pred. No. 3.1e-127;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VELQLSRVNHPIVVKLYGACINPVCLVMEYAEAGGSLYNVLHGABPLPYTAAHAMSWCL 60  
DB 76 VELQLSRVNHPIVVKLYGACINPVCLVMEYAEAGGSLYNVLHGABPLPYTAAHAMSWCL 135  
QY 61 QCSQGVAVLHSMQPKALHRLDKPPNLLVAGGTVLKICDFGTACDIOTHTMTNKGSAW 120  
DB 136 QCSQGVAVLHSMQPKALHRLDKPPNLLVAGGTVLKICDFGTACDIOTHTMTNKGSAW 195  
QY 121 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWVHNGTRPPLIKNL 180  
DB 196 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWVHNGTRPPLIKNL 255  
QY 181 PKPIESLMTRCWSKDPSPQSPMSMEIIVKIMTHLMRYFPGADEPLQYPCQ 228  
DB 256 PKPIESLMTRCWSKDPSPQSPMSMEIIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 8  
AA09547  
ID AAY09547 standard; Protein; 590 AA.  
XX  
AC AAY09547;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE Human TAK1-6xHis protein.  
XX  
KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
KW transforming growth factor beta.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO921010-A1.  
XX  
PD 29-APR-1999.  
XX  
PF 22-OCT-1998; 98WO-JP04796.  
XX  
PR 22-OCT-1997; 97JP-0290188.  
XX  
PA (CHUS) CHUGAI SEIYAKU KK.  
XX  
PI Ohtomo T, Ono K, Tsuchiya M;  
XX  
DR WPI; 1999-312645/26.  
DR N-PSDB; AAX56285.  
XX  
PT Screening for TGF- beta inhibitory substances, which are useful as  
PT drugs for treatment of diseases relating to its disorder  
XX  
PS Example 1; Page 171-174; 195pp; Japanese.

A method has been developed for screening for substances which inhibit the binding of TAK1 polypeptide to TAB1 polypeptide. The method comprises: (a) contacting the polypeptide in the presence of a sample; and (b) detecting the amount of bound polypeptide, in which the sample

can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming growth factor (TGF)-beta inhibitory substances can be used in drugs for indications e.g. as TGF-beta signal transmission inhibitors or activators, or extracellular matrix protein production enhancement inhibitors, or activators, or cell proliferation prevention inhibitors or activators, or monocyte migration inhibitors or activators, or physiological activity induction inhibitors or activators, or immunosuppression inhibitors or activators, or amyloid beta protein precipitation inhibitors or activators, and such substances can also be inhibitors of the TAK1 polypeptide function, particularly kinase activity. The present sequence represents TAK1-6xHis from an example of the present invention.

Sequence 590 AA;  
Query Match 100.0%; Score 1252; DB 20; Length 590;  
Best Local Similarity 100.0%; Pred. No. 3.2e-127;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VELQLSRVNHPIVVKLYGACINPVCLVMEYAEAGGSLYNVLHGABPLPYTAAHAMSWCL 60  
DB 76 VELQLSRVNHPIVVKLYGACINPVCLVMEYAEAGGSLYNVLHGABPLPYTAAHAMSWCL 135  
QY 61 QCSQGVAVLHSMQPKALHRLDKPPNLLVAGGTVLKICDFGTACDIOTHTMTNKGSAW 120  
DB 136 QCSQGVAVLHSMQPKALHRLDKPPNLLVAGGTVLKICDFGTACDIOTHTMTNKGSAW 195  
QY 121 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWVHNGTRPPLIKNL 180  
DB 196 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWVHNGTRPPLIKNL 255  
QY 181 PKPIESLMTRCWSKDPSPQSPMSMEIIVKIMTHLMRYFPGADEPLQYPCQ 228  
DB 256 PKPIESLMTRCWSKDPSPQSPMSMEIIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 9  
AAY28997  
ID AAY28997 standard; Protein; 606 AA.  
XX  
AC AAY28997;  
XX  
DT 29-OCT-1999 (first entry)  
XX  
DE Human TGF-beta activated kinase (TAK) 1b amino acid sequence.  
XX  
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;  
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;  
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;  
KW endotoxin shock; septicemia; human; HTAK1b.  
XX  
OS Homo sapiens.  
XX  
PN WO9940202-A1.  
XX  
PD 12-AUG-1999.  
XX  
PF 02-FEB-1999; 99WO-JP00422.  
XX  
PR 30-OCT-1998; 98JP-0309316.  
PR 06-FEB-1998; 98JP-0026003.  
XX  
PA (TANA) TANABE SEIYAKU CO.  
XX  
PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;  
XX  
DR WPI; 1999-494298/41.  
DR N-PSDB; AAX99697.  
XX  
PT Nuclear factor kappa B activation inhibitors, useful as preventives  
XX for, e.g. autoimmune diseases  
PS Examples; Page 39-43; 49pp; Japanese.

XX The invention provides a method for identifying or screening a nuclear  
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a  
 CC test substance on modulating the function(s) of TGF-beta activated kinase  
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
 CC intractable diseases with inflammation (such as atrophic dermatitis and  
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The  
 CC present sequence represents the amino acid sequence of human TAK1b  
 CC (hTAK1b) protein.

XX SQ Sequence 606 AA;  
 Query Match 100.0%; Score 1252; DB 20; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-127;  
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 VELRQLSRVNHPIVLYGACLPVCLVMEYAEAGGSLYNVLHGAEPLPYTTAAHMSWCL 60  
 Db 76 VELRQLSRVNHPIVLYGACLPVCLVMEYAEAGGSLYNVLHGAEPLPYTTAAHMSWCL 135  
 QY 61 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPGTACDIQTHMTNKGSAW 120  
 Db 136 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPGTACDIQTHMTNKGSAW 195  
 QY 121 MAPEVPEGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180  
 Db 196 MAPEVPEGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255  
 QY 181 PKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLQPCQ 228  
 Db 256 PKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLQPCQ 303

## RESULT 10

ABBS8061  
 ID ABB58061 standard; Protein; 678 AA.  
 XX AC ABB58061;  
 XX DT 26-MAR-2002 (first entry)  
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 975.  
 XX DE Drosophila; developmental biology; cell signalling; insecticide;  
 XX KW pharmaceutical.  
 XX OS Drosophila melanogaster.  
 XX FN WC200171042-A2.  
 XX PD 27-SEP-2001.  
 XX PF 23-MAR-2001; 2001WO-US09231.  
 XX PR 23-MAR-2000; 2000US-191637P.  
 XX PR 11-JUL-2000; 2000US-0614150.  
 XX XX (PEKE ) PE CORP NY.  
 XX XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 XX DR N-PSDB; ABL02164.  
 XX XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 FT genes from Drosophila and for elucidating cell signalling and cell-cell  
 FT interactions -  
 XX PS Disclosure; SEQ ID NO 975; 21pp + Sequence Listing; English.  
 XX CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at fip.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 678 AA;  
 Query Match 55.9%; Score 700; DB 22; Length 678;  
 Best Local Similarity 57.5%; Pred. No. 5.4e-67;  
 Matches 130; Conservative 35; Mismatches 57; Indels 4; Gaps 3;  
 QY 2 ELRQLSRVNHPIVLYG--ACLPVCLVMEYAEAGGSLYNVLHGAEPLPYTTAAHMSWCL 59  
 Db 60 EVKQLSRVNHPIVLYG--ACLPVCLVMEYAEAGGSLYNVLHGAEPLPYTTAAHMSWCL 118  
 QY 60 LQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPGTACDIQTHMTNKGSAW 119  
 Db 119 RQCAEGLAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPGTACDIQTHMTNKGSAW 178  
 QY 120 WMAPEVPEGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKN 179  
 Db 179 WMAPEVPEGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKN 237  
 QY 180 LKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLQPCQ 225  
 Db 238 CPKRIEDLMTACWKVPEDEPSMQYIVGVHMEIVKDYTGADKALEY 283

## RESULT 11

ABBS60985  
 ID ABB60985 standard; Protein; 252 AA.  
 XX AC ABB60985;  
 XX DT 26-MAR-2002 (first entry)  
 XX DE Drosophila melanogaster polypeptide SEQ ID NO 9747.  
 XX DE Drosophila; developmental biology; cell signalling; insecticide;  
 XX KW pharmaceutical.  
 XX OS Drosophila melanogaster.  
 XX FN WC200171042-A2.  
 XX PD 27-SEP-2001.  
 XX PF 23-MAR-2001; 2001WO-US09231.  
 XX PR 23-MAR-2000; 2000US-191637P.  
 XX PR 11-JUL-2000; 2000US-0614150.  
 XX XX (PEKE ) PE CORP NY.  
 XX XX Venter JC, Adams M, Li PWD, Myers EW;  
 XX WPI; 2001-656860/75.  
 XX DR N-PSDB; ABL05089.  
 XX XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 FT genes from Drosophila and for elucidating cell signalling and cell-cell  
 FT interactions -  
 XX PS Disclosure; SEQ ID NO 9747; 21pp + Sequence Listing; English.  
 XX CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and



cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 252 AA;

Query Match 33.1%; Score 414; DB 22; Length 252;  
Best Local Similarity 41.8%; Pred. No. 1.9e-36;  
Matches 82; Conservative 35; Mismatches 53; Indels 26; Gaps 4;

QY 2 ELRQLSRVNHENIVKLYGACLNIVC--LVMEVAGSGSYLVHGAEPDLPYTAHAMSVC 59  
Db 54 EYQUTKASHVNIIVELYGTSRHEGCALLMEFVGGSLSSFLH-AKSKPSYSHAHAFNWA 112  
QY 60 LQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTGLKICDFGTACDIQTMTNNKGSAA 119  
Db 113 HQIAGIAYLHGMQPKAVIHRDIKPEINTLLCEKGLKICDPTVVDLSQSTSCNAGTCR 172  
QY 120 WMAPE-----VFEGSNYSEKCDVFSWGILMEVITRRKPPDEI 157  
Db 173 YKAPEVRELEDFKSNRIINOPTGFKVLQGNKPKDCKDVYSWALTFWELLSRKEPFQY 232  
QY 158 GGPAPRIMWAVHNGTR 173  
Db 233 -NTLFELYMAINEGR 247

RESULT 12  
AAG32053  
ID AAG32053 standard; Protein; 367 AA.  
AC AAG32053;  
DT 17-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38597.  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38597.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
PR 05-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.



PR	28-APR-1999;	99US-0131449.	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	99US-0132048.	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	99US-0132407.	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	99US-0132484.	PR	22-JUL-1999;	99US-0145087.
PR	05-MAY-1999;	99US-0132485.	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145192.
PR	06-MAY-1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145145.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145218.
PR	11-MAY-1999;	99US-0134256.	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	99US-0134218.	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145918.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145919.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145951.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0146386.
PR	20-MAY-1999;	99US-0135124.	PR	02-AUG-1999;	99US-0146388.
PR	21-MAY-1999;	99US-0135353.	PR	02-AUG-1999;	99US-0146389.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0147038.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147204.
PR	27-MAY-1999;	99US-0136392.	PR	04-AUG-1999;	99US-0147302.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	99US-0137222.	PR	05-AUG-1999;	99US-0147152.
PR	03-JUN-1999;	99US-0137528.	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	99US-0137724.	PR	06-AUG-1999;	99US-0147303.
PR	08-JUN-1999;	99US-0138094.	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	99US-0138540.	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	99US-0138847.	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	99US-0139119.	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	99US-0139452.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139453.	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139750.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139763.	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	99US-0139817.	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	99US-0139899.	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144089.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.

```

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.

Query Match 29.98; Score 374; DB 21; Length 369;
Best Local Similarity 39.48; Pred. No. 7.4e-32;
Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHPIVVKLYGACINPV--CLVMEYAEGLSYNVL---HGAEPLPYTTAAHA 55
Db 137 EVSMLANLKHPIVRFVIGACKPMVWCIVTEYAKGVSVRQFLTRQNRVPLKL-----A 191

QY 56 MSWCLOCSQGVAVLHNSQPKALIHRLDKPNNLLVAGTVLKI CDFGTA-CDIQTH-MTN 113
Db 192 VKOALDVARGMAYVHG---RNFIHRLDKSNLLISADKSI-KIADFGVARIEVQTEGMTP 247

QY 114 NKGSAAWMAPEVFEKSGNSYSEKCVFSGIILWEIVTERKPEDEIGG--PAFRIMWVHNG 171
Db 248 ETGYRWMAPEMIQHRAYNOKVDYISGIVLWELITGLLPFQNNTVQRAFAV---VNRG 304

QY 172 TRPLIKNLKPIESLMTRCWSDKPSQPSMEEIVKIM 209
Db 305 VRTVNDCLPVLSDIMTRCDANPEVRPCFVEVVKLL 342

RESULT 14
AAG32051
ID AAG32051 standard; Protein; 407 AA.
AC AAG32051;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38595.
DE
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127452.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135829.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

```

PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159337.
PR 21-JUL-1999;	99US-0144884.	PR 18-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 28-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147302.	PR 29-OCT-1999;	99US-0162142.
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148664.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149375.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149723.		
PR 23-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		
PR 23-SEP-1999;	99US-0155486.		
PR 24-SEP-1999;	99US-0155659.		
PR 28-SEP-1999;	99US-0156458.		
PR 29-SEP-1999;	99US-0156596.		
PR 04-OCT-1999;	99US-0157117.		
PR 05-OCT-1999;	99US-0157753.		
PR 06-OCT-1999;	99US-0157865.		
PR 07-OCT-1999;	99US-0158029.		
PR 08-OCT-1999;	99US-0158232.		
PR 12-OCT-1999;	99US-0158369.		
PR 13-OCT-1999;	99US-0159293.		
PR 13-OCT-1999;	99US-0159294.		
PR 13-OCT-1999;	99US-0159295.		

Query Match 29.9%; Score 374; DB 21; Length 407;

Best Local Similarity 39.4%; Pred. No. 8.6e-32;

Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

Qy	2	ELROLSRVNHPNIVKLYGACLNVP--CLVMEYAEAGGSLXNVL---HGAELPLPYTAHA 55
Db	175	EVSMLANLKHPNIVRFICACRKPWVCIVTEYAKGSGVRQFLTRQNRVPLKL-----A 229
Qy	56	MSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTIVLKICDFGA-CDIOTH-MTN 113
Db	230	VKQALDVARGWAYVHG--RNFIHRLDKSDNLLISADKSI-KIADFGVARIEVQTEGMTP 285
Qy	114	NGSAAMWAEVPEGSNYSEKDFVSGIILWEVITRKPDEIGG--PAPRIMWYHNG 171
Db	286	ETGYRWMAPEMIQHRAYNOKVDYSEFIVLWELITGLLPFQNNMTAVQAAFAV---VNRG 342
Qy	172	TRPLIKNLKPTESLMTRCNSKDPORSPSEEVVKIM 209
Db	343	VRPIVNDCLFVLSDIMTRCNDANPEVRPCFEVVKLL 380

RESULT 15

AAG22172  
ID AAG22172 standard; Protein; 369 AA.

XX AC AAG22172;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 24999.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.

XX OS Arabidopsis thaliana.

XX FN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-01322407.  
PR 04-MAY-1999; 99US-01322484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140895.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0143624.  
PR 16-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 22-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145313.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.

```
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      29.8%; Score 373; DB 21; Length 369;
Best Local Similarity 39.9%; Pred No. 9.5e-32;
Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;

Qy 2 ELRQLSRVNHENIVKLYGACLNPV--CLVMEYAEKGSGLYNNVL---HGAELPLPYTAAHA 55
Db 137 EVSMLAEFKHENIVRFICGACIKPMWCIVTEYAKGSVYRQFLTKRQNEAVPLK-----A 191

Qy 56 MSWCLOCSQGVAYLHSMOPKALIHRLKPPNLLVAGGTVLKICDFGTA-CDIQTH-MTN 113
Db 192 VMQALDVARGMAYVHE---RNFIHRDLKSDNLLISADRSI-KIADFGVARIEVQTEGMP 247

Qy 114 NKGSAAWAPRVEFGSNYSKCDVFSWGIILWEVILTRKRPDEIGG--PAPRIMWAVHNG 171
Db 248 ETGYRWAPENIQHRTYQVDVYSFGIVLWELITGLLPQNMTAVQAFAV---VNRG 304

Qy 172 TRPFLIKNLPKPIESLMTRCWSKDPSPQSPSEIIVKIM 209
Db 305 VRPTVPADCLPVLGEIMTRCWDADPEVPCFAEIVNLL 342
```

Search completed: December 4, 2003, 09:30:08  
Job time : 78.7973 secs

This Page Blank (18pt)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:21:16 ; Search time 95.5135 Seconds  
(without alignments)  
615.997 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303

Perfect score: 1252

Sequence: 1 VELQLSRVNHPIVKLYGA.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_virus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1252	100.0	478	4	Q9NTR4	Q9ntr4 homo sapien
2	1252	100.0	491	4	Q9NZ70	Q9nz70 homo sapien
3	1252	100.0	539	4	Q9NTR1	Q9ntr1 homo sapien
4	1252	100.0	566	4	Q9NTR2	Q9ntr2 homo sapien
5	1252	100.0	606	11	Q923A8	Q923a8 mus musculus
6	1221	97.5	616	13	Q73613	O73613 xenopus lae
7	700	55.9	678	5	Q9V3Q6	Q9v3q6 drosophila
8	414	33.1	258	5	Q9VCV0	Q9vcv0 drosophila
9	379	30.3	422	10	Q8GV29	Q8gv29 oryza sativ
10	376	30.0	371	13	Q90ZV8	Q90zv8 brachydanio
11	375.5	30.0	2631	5	Q8MVR1	Q8mvr1 dictyosteli
12	374.5	29.9	2964	5	Q817W7	Q817w7 dictyosteli
13	374	29.9	411	10	Q9ZQ31	Q9zq31 arabidopsis
14	373	29.8	412	10	Q9M085	Q9m085 arabidopsis
15	373	29.8	417	10	Q8GV30	Q8gv30 oryza sativ
16	372	29.7	289	11	Q8BR73	Q8br73 mus musculus

17	372	29.7	454	11	Q9ESL3	Q9esl3 mus musculus
18	372	29.7	802	11	Q9ESL4	Q9esl4 mus musculus
19	371	29.6	455	4	Q9HCC4	Q9hcc4 homo sapien
20	371	29.6	800	4	Q9NYE9	Q9nye9 homo sapien
21	371	29.6	800	4	Q9NYL2	Q9nyl2 homo sapien
22	371	29.6	800	4	Q9HCC5	Q9hcc5 homo sapien
23	371	29.6	800	4	Q9HDD2	Q9hdd2 homo sapien
24	368	29.4	416	10	Q94C42	Q94c42 triticum ae
25	366	29.2	637	10	Q94J41	Q94j41 oryza sativ
26	363.5	29.0	1148	5	Q95VF6	Q95vf6 drosophila
27	363.5	29.0	1161	5	Q95UN8	Q95un8 drosophila
28	362	28.9	422	5	Q23846	Q23846 dictyosteli
29	362	28.9	1338	5	Q23927	Q23927 dictyosteli
30	360.5	28.8	608	11	Q8B1G8	Q8b1g8 mus musculus
31	360.5	28.8	1066	4	Q9H2N5	Q9h2n5 homo sapien
32	356.5	28.5	888	11	Q8CBX3	Q8cbx3 mus musculus
33	356	28.4	462	10	Q39886	Q39886 glycine max
34	355.5	28.4	370	10	Q9S7D5	Q9s7d5 arabidopsis
35	355	28.4	421	10	Q8GV28	Q8gv28 oryza sativ
36	355	28.4	564	4	Q9HLY7	Q9hly7 homo sapien
37	354.5	28.3	1161	5	Q8MRK7	Q8mrk7 drosophila
38	353.5	28.2	859	4	Q8WY25	Q8wy25 homo sapien
39	353.5	28.2	888	11	Q8CDL6	Q8cdl6 mus musculus
40	351.5	28.1	977	5	Q9VM24	Q9vm24 drosophila
41	351	28.0	570	4	Q8WMN2	Q8wmn2 homo sapien
42	351	28.0	1036	4	Q8WWN1	Q8wwn1 homo sapien
43	350	28.0	546	10	O22558	O22558 arabidopsis
44	349	27.9	1001	11	Q8VDG6	Q8vdg6 mus musculus
45	348	27.8	391	10	Q8LCP3	Q8lcp3 arabidopsis

#### ALIGNMENTS

#### RESULT 1

Q9NTR4 PRELIMINARY; PRT; 478 AA.  
AC Q9NTR4  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE DJ154G14.1.3 (Mitogen-activated protein kinase kinase 7 (TGF-beta activated kinase 1c (TAK1))) (Fragment).  
GN MAP3K7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]\_TaxID=9606;  
RP SEQUENCE FROM N.A.  
RA Tracey A.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AL121964; CAB87604.1; --  
DR HSP; P12931; IFMK.  
DR InterPro; IPR000719; Prot\_kinase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
FT NON TER 1  
SQ SEQUENCE 478 AA; 52482 MW; 177C8CFA8D8DBF8 CRC64;

Query Match 100.0%; Score 1252; DB 4; Length 478;  
Best Local Similarity 100.0%; Pred. No. 3.4e-121;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELQLSRVNHPIVKLYGACLNIPVCLVMEYAEGLSYNLVHGAEPDLPYTAHAMSWCL 60

```

Db 36 VELRQLSRVNHPTIVKLYGACLVCLVMEYAEAGSLYNVLHGAEPLPYTAAHMSWCL 95
Qy 61 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTGACDIQTHMTNNKGSAAW 120
Db 96 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTGACDIQTHMTNNKGSAAW 155
Qy 121 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
Db 156 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 215
Qy 181 PKPIESLMTRCWSKDPSPQSRPSMEIEIVKIMTHLMRYFFGADPELPQPCQ 228
Db 216 PKPIESLMTRCWSKDPSPQSRPSMEIEIVKIMTHLMRYFFGADPELPQPCQ 263

RESULT 2
Q9NZ70 Q9NZ70 PRELIMINARY; PRT; 491 AA.
AC Q9NZ70; Q9NZ70;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TGF beta-activated kinase splice variant d (DJ154G14.1.4) (Mitogen-
activated protein kinase kinase 7 (TGF-beta activated kinase 1d
(TAK1)))
DE (TAK1))
GN TAK1 OR MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568765; PubMed=11118615;
RA Dempsey C.E., Sakurai H., Sugita T., Guesdon F.;
RT "Alternative splicing and gene structure of the transforming growth
factor beta-activated kinase 1."
RT factor beta-activated kinase 1."
RL Biochim. Biophys. Acta 1517:46-52(2000).
RN [2]
RP SEQUENCE OF 41-491 FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF218074; AAF27652.1; -.
DR EMBL; AL218074; AAF27652.1; -.
DR HSP; P08631; IAD5.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR001245; Ser thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER 1
FT SEQUENCE 491 AA; 53739 MW; BYD8832E286A99CS CRC64;

Query Match 100.0%; Score 1252; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 3.5e-121;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VELRQLSRVNHPTIVKLYGACLVCLVMEYAEAGSLYNVLHGAEPLPYTAAHMSWCL 60
Db 76 VELRQLSRVNHPTIVKLYGACLVCLVMEYAEAGSLYNVLHGAEPLPYTAAHMSWCL 135
Qy 61 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTGACDIQTHMTNNKGSAAW 120
Db 136 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTGACDIQTHMTNNKGSAAW 195
Qy 121 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
Db 196 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255

```

```

Qy 181 PKPIESLMTRCWSKDPSPQSRPSMEIEIVKIMTHLMRYFFGADPELPQPCQ 228
Db 256 PKPIESLMTRCWSKDPSPQSRPSMEIEIVKIMTHLMRYFFGADPELPQPCQ 303

RESULT 3
Q9NTR1 Q9NTR1 PRELIMINARY; PRT; 539 AA.
AC Q9NTR1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DJ154G14.1.1 (Mitogen-activated protein kinase kinase 7 (TGF-
beta activated kinase 1a (TAK1))) (Fragment).
DE beta activated kinase 1a (TAK1)) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL218074; CAB87607.1; -.
DR HSP; P12931; IFMK.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON TER 1
FT SEQUENCE 539 AA; 60006 MW; E6183F553CC7F324 CRC64;

Query Match 100.0%; Score 1252; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 4e-121;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VELRQLSRVNHPTIVKLYGACLVCLVMEYAEAGSLYNVLHGAEPLPYTAAHMSWCL 60
Db 36 VELRQLSRVNHPTIVKLYGACLVCLVMEYAEAGSLYNVLHGAEPLPYTAAHMSWCL 95
Qy 61 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTGACDIQTHMTNNKGSAAW 120
Db 96 QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKI CDFTGACDIQTHMTNNKGSAAW 155
Qy 121 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
Db 156 MAPEVFEAGSYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 215
Qy 181 PKPIESLMTRCWSKDPSPQSRPSMEIEIVKIMTHLMRYFFGADPELPQPCQ 228
Db 216 PKPIESLMTRCWSKDPSPQSRPSMEIEIVKIMTHLMRYFFGADPELPQPCQ 263

RESULT 4
Q9NTR2 Q9NTR2 PRELIMINARY; PRT; 566 AA.
AC Q9NTR2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DJ154G14.1.2 (Mitogen-activated protein kinase kinase 7 (TGF-
beta activated kinase 1b (TAK1))) (Fragment).
DE beta activated kinase 1b (TAK1)) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```





RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Foubler K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupeki M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassaman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragov V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirkas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of *Drosophila melanogaster* genome";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of *Drosophila melanogaster* genome";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003740; AAF56055.2; -  
DR FlyBase; FBgn039015; Tak12.  
SQ SEQUENCE 258 AA; 29480 MW; D117DA8922BF8AE1E CRC64;  
Query Match 33.1%; Score 414; DB 5; Length 258;  
Best Local Similarity 41.8%; Pred. No. 1.2e-34;  
Matches 82; Conservative 35; Mismatches 53; Indels 26; Gaps 4;  
QY 2 ELRQLSRVHNHNIVKLYGACLNPCV--LVMEYAGSGSLYVNLHGAEPLPYVTAHAMSWC 59

Db 54 EYQLTAKSHVNIYELYTSRHEGACALLMEFVDCGSLSSFLH-AKSPSYSHAHAFNWA 112  
QY 60 LQCSQGVAYLHSMQKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMNNKGSAA 119  
Db 113 HQIAQGIAYLHGMQPKAVIHRDIKPLNTLLCEKGLKICDFGTVDLSQSISCNAGTCR 172  
QY 120 WMADE-----VFESNYSYSEKCVFSWGIIILWEVITRRKPEDEIGG--PAFRIMVAVHNG 157  
Db 173 YKAPEVRELFDFKGNRIINOFTGQKVLQGNKPKDEKCDVYSWAIPTFWELLSRKPEFQY 232  
QY 158 GGPAPRIMVAVHNGTR 173  
Db 233 -NTLPELYMAINEGKR 247  
RESULT 9  
Q8GV29 PRELIMINARY; PRT; 422 AA.  
AC Q8GV29  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Serine/threonine protein kinase.  
GN DPX2.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gu Z.M., Zhang H.S., Huang J.;  
RT "Molecular Cloning And Characterization Of A Novel Serine/Threonine  
RT Protein Kinase Gene Family From Rice."  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY156511; AAN84503.1; -  
KW Kinase.  
SQ SEQUENCE 422 AA; 46587 MW; BCC22F1FC6CF62C CRC64;  
Query Match 30.3%; Score 379; DB 10; Length 422;  
Best Local Similarity 38.6%; Pred. No. 8.9e-31;  
Matches 85; Conservative 43; Mismatches 70; Indels 22; Gaps 9;  
QY 2 ELRQLSRVHNHNIVKLYGACLNPCV--CLVMEYAGSGSLYVNL----HGAELPYVTAHA 55  
Db 190 EVMLATLIRHENIVKFIAGCRKPMVWCIVTEYAKGSGVRQFLMKRQNSVPLKL-----A 244  
QY 56 MSWCLQCSQGVAYLHSMQKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTH-MTN 113  
Db 245 VKQALDVARGWAYVHAL---GFIHRDLKSDN-LIISGDKSIKIADFGVARIEVTEGTP 300  
QY 114 NGSAAWMAPEVRFESNYSYSEKCVFSWGIIILWEVITRRKPEDEIGG--PAFRIMVAVHNG 171  
Db 301 ETGYRWMAPEMIQHPYDQKVDVYSGIVLWELITGLMPLFANMTVAQAFAV---VNGK 357  
QY 172 TRPLIKNLKPKIESLMTKRCWSDPSQPSMEEIVKIMTH 211  
Db 358 VRPAIPQDCLPVLSEIMTRCWDPNPDVPRPPPTTEVVRMLEH 397  
RESULT 10  
Q90ZY8 PRELIMINARY; PRT; 371 AA.  
ID Q90ZY8  
AC Q90ZY8  
DT 01-DEC-2001 (TRENBLrel. 19, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Protein kinase Npk.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.

```

OX NCBI_TaxID=7955;
RN
RP SEQUENCE FROM N.A.
RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;
RT "A novel protein kinase, zNPK, from the zebrafish.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF265343; AA852416.1; -
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ
SEQUENCE 371 AA; 42456 MW; 989188A8B20D296 CRC64;

Query Match 30.0%; Score 376; DB 13; Length 371;
Best Local Similarity 39.0%; Pred. No. 1.6e-30;
Matches 82; Conservative 37; Mismatches 81; Indels 10; Gaps 7;

QY 2 ELRQLSRVHPNIVKLYGACLNVPV-CLVMEYAGGSLYVNLHGAEPLPYTAHAHMSWC 59
DB 84 EAILSVLSHKNITQFYGAILEAPNDGIVTEYASRGSLEYLSSADS-EEMDMQVMTWA 142

QY 60 LQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTGLKICDGFACDIQTHMTNKK--GS 117
DB 143 MEIAGKWHYLAHAAPLVHRLDKSRNVVLTVA-DNVLKICDFG-ASKWVSHHTMSLVGT 200

QY 118 AAMWAPVFEVGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPL 176
DB 201 FFWNAPEVIQSLPVSEICDYSYGVLMWMLTREVPFK--GFEGLOVAVLWVEXHERPTI 258

QY 177 IKNLKPPIESIMTRCWSKQSPSMBEIV 206
DB 259 PSSCPASFADLMRCWNAEPKRPQFKQIL 288

RESULT 11
Q8MVRI PRELIMINARY; PRT; 2631 AA.
AC Q8MVRI;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cyclic GMP-binding protein C.
GN GBPC.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN
RP SEQUENCE FROM N.A.
RA Bosgraaf L., Ruescher H., Smith J.L., Wessels D., Solls D.R.,
RA Van Haastert P.J.M.;
RT "A novel cGMP signalling pathway mediating myosin phosphorylation and
RT chemotaxis in Dictyostelium.";
RL EMBL J. 0:0-0(2002).
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=22008080; PubMed=12011437;
RA Goldberg J.M., Bosgraaf L., Van Haastert P.J., Smith J.L.;
RT "Identification of four candidate cGMP targets in Dictyostelium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:6749-6754(2002).
DR EMBL: AF481923; AAM34041.1; -
DR InterPro: IPR000595; cNMP_Binding.
DR InterPro: IPR000591; DEP.
DR InterPro: IPR004182; GRAM_dom.
DR InterPro: IPR003591; LRR_typ.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR000651; RasGEF.
DR InterPro: IPR001895; RasGEF_CDC25.
DR InterPro: IPR001806; Ras_trnsfrmg.
DR InterPro: IPR002290; Ser_thr_pkinase.

```

```

DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00027; cNMP_binding; 2.
DR Pfam: PF00610; DEP; 1.
DR Pfam: PF02893; GRAM; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00617; RasGEF; 1.
DR Pfam: PF00618; RasGEFN; 1.
DR PRINTS: PR00449; RASTRNSFRMG.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00049; DEP; 1.
DR SMART: SM00568; GRAM; 1.
DR SMART: SM00369; LRR_TYP; 2.
DR SMART: SM00147; RasGEF; 1.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TYK; 1.
DR PROSITE: PS00042; cNMP_BINDING_3; 2.
DR PROSITE: PS00186; DEP; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; GTP-binding; Transferase.
SQ
SEQUENCE 2631 AA; 294124 MW; 5857AABDOA637701 CRC64;

Query Match 30.0%; Score 375.5; DB 5; Length 2631;
Best Local Similarity 35.8%; Pred. No. 1.9e-29;
Matches 88; Conservative 44; Mismatches 75; Indels 39; Gaps 9;

QY 2 ELRQLSRVHPNIVKLYGACLNVPVCLVMEYAGGSLYVNLHGAEPLPYTAHAHMSWCL- 60
DB 947 EVYMSGNLHPNVMKISGFCIQPLCMALYVRYGSLYSLLS-----NSSIEISWGLR 998

QY 61 -----QCSQGVAYLHSMQPKALIHRLDKPPNLL--VAGG--TVLKI CDGCTACDIQTHM 111
DB 999 LQIASEIAKQWHLHSHNP-FVTHRLDKSPNILLINGITEGNSVATIDFGTS----- 1050

QY 112 TNKNGSAA-----WMAPEVFEVGSNYSEKCDVFSWGIILWEVITRRKPFDE--IGGP 160
DB 1051 TALYGGRAALRCVDQPLWLGPEVLGATAYSEPSDVYSFGIILWLYTRAHPFDEFQFGQW 1110

QY 161 AFRIMWAVHNGTRPPLIKNLKPPIESIMTRCWSKQSPSMBEIVKIMTHLMRYFPQAD 220
DB 1111 MSKLEDEIIRGLRPTPTPTCPPEYVELIQSCWTHPEPSRPTFTSIVEILGQIKKFF--AP 1168

QY 221 EPLQYP 226
DB 1169 LPFTHP 1174

RESULT 12
Q817W7 PRELIMINARY; PRT; 2964 AA.
AC Q817W7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Pats1 (Fragment).
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN
RP SEQUENCE FROM N.A.
RA Aysaiah J.C., Kuchnick L.L., Laroche D.A.;
RT "The identification of Pats1, a novel gene locus required for
RT cytokinesis in Dictyostelium discoideum.";
RL Mol. Biol. Cell 0:0-0(2003).
DR EMBL: AY170918; AAO12857.1; -
FT NON TER 1
FT NON TER 2964 2964
SQ
SEQUENCE 2964 AA; 331830 MW; 0905BC2137318AB2 CRC64;

Query Match 29.9%; Score 374.5; DB 5; Length 2964;
Best Local Similarity 40.1%; Pred. No. 2.8e-29;

```



RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
RT "Arabidopsis cDNA clones";  
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
[6]  
RN SEQUENCE FROM N.A.  
RP Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,  
RA Chan M.M., Chang C.H., Dale J.M., Deng J.M., Hayashizaki Y.,  
RA Hsuan V.W., Ishida J., Jones T., Kaniya A., Karlin-Neumann G.,  
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
RA Nguyen M., Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M.,  
RA Southwick A., Tang C.C., Toriumi M., Wallender E.K., Wong C., Wu H.C.,  
RA Yamada K., Yu G., Yuan S., Shinozaki K., Davis R.W., Theologis A.,  
RA Ecker J.R.;  
RT "Arabidopsis ORF clones";  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AL161578; CAB79835.1; -;  
DR EMBL; AY085535; AAM62759.1; -;  
DR EMBL; AY125513; AAM78105.1; -;  
DR EMBL; BT000587; AAN18156.1; -;  
DR HSP; P08631; IAD5.  
DR InterPro; IPR000719; Prot.kinase.  
DR InterPro; IPR002290; Ser.thr.pkinase.  
DR InterPro; IPR001245; Tyr.pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot.kinase; 1.  
DR PROSITE; PS00011; PROTEIN KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
DR KATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 412 AA; 46083 MW; 644F35A90210D488 CRC64;  
  
Query Match 29.8%; Score 373; DB 10; Length 412;  
Best Local Similarity 39.9%; Pred. No. 3.6e-30;  
Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;  
  
QY 2 ELRQLSRVNHENIVKLYGACLNVP--CLVMEYAEAGGSLYNVL-----HGAELPLPYTAHA 55  
DB 180 EVSMLAFKHNIVRFVIGACIKPWCWICIVTEYAKGSGVRQFLTKRQNFVPLKL-----A 234  
  
QY 56 MSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTA-CDIOTH-MTN 113  
DB 235 VMQALDVARGMAYVHVE---RNFIHRLDKSDNLLISADRSI-KIADFGVARIEVTEGMTP 290  
  
QY 114 NKGSAAMWAPVEFGSNYSKCDVFSWGIILWEVITRKPFDEIGG--PAFRIMWYHNG 171  
DB 291 ETGTYRWMAPEVIOHRPYTKQVDVYSGFVILWELITGLLPQNTAVQAFAV---VNRG 347  
  
QY 172 TRPLIKNLKPPIESLMTRCWSKDPSPORSMEIEIVKIM 209  
DB 348 VRPTVPADCLPVIGELMTRCWDADPEVRPCFAEIVNLL 385  
  
RESULT 15  
Q8GV30 PRELIMINARY; PRT; 417 AA.  
AC Q8GV30;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Serine/threonine protein kinase.  
GN DP1.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=39947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gu Z.M., Zhang H.S., Huang J.;  
RT "Molecular Cloning And Characterization Of A Novel Serine/Threonine  
RT Protein Kinase Gene Family From Rice";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY156510; AAN84502.1; -;  
KW Kinase.  
SQ SEQUENCE 417 AA; 45927 MW; 1B66E3596653FAD7 CRC64;  
  
Query Match 29.8%; Score 373; DB 10; Length 417;  
Best Local Similarity 38.5%; Pred. No. 3.7e-30;  
Matches 84; Conservative 41; Mismatches 71; Indels 22; Gaps 9;  
  
QY 2 ELRQLSRVNHENIVKLYGACLNVP--CLVMEYAEAGGSLYNVLHGAE-----PLPYTAHA 55  
DB 185 EVSMLATLRHSNIVKFGACRKPWCWICIVTEYAKGSGVRNRLRQNRSPVPLKL-----A 239  
  
QY 56 MSWCLOCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTA-CDIOTH-MTN 113  
DB 240 VKQALDVARGMAYVHGL---GFHRLDKSDN-LIISGDKIKIADFGVARIEVTEGMTP 295  
  
QY 114 NKGSAAMWAPVEFGSNYSKCDVFSWGIILWEVITRKPFDEIGG--PAFRIMWYHNG 171  
DB 296 ETGTYRWMAPEVIOHRPYTKQVDVYSGFVILWELITGLLPQNTAVQAFAV---VNRG 352  
  
QY 172 TRPLIKNLKPPIESLMTRCWSKDPSPORSMEIEIVKIM 209  
DB 353 VRPAIPHDCLPALAEIMTRCWDANPDARPPFTEVVRML 390  
  
Search completed: December 4, 2003, 09:33:02  
Job time : 96.5135 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 09:34:57 ; Search time 4908.16 Seconds  
(without alignments)  
1900.384 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHPIVVKLYGA.....MTHLMRYPPGADEPLQYPCQ 228

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09830144/runat\_03122003\_122509\_23321/app\_query.fasta\_1.654  
-DB=GenEmbl -QWMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144@cgn\_1\_1\_3608 @runat\_03122003\_122509\_23321 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.to:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.or:\*  
21: em.on:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.to:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vi:\*  
30: em.htg\_hum:\*  
31: em.htg\_inv:\*  
32: em.htg\_other:\*  
33: em.htg\_mus:\*  
34: em.htg\_pln:\*  
35: em.htg\_rod:\*  
36: em.htg\_mam:\*  
37: em.htg\_vrt:\*  
38: em.sy:\*  
39: em.htgo\_hum:\*  
40: em.htgo\_mus:\*  
41: em.htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1252	100.0	1704	6	E38399	E38399 NF-kappa B
2	1252	100.0	1705	9	AB009358	AB009358 Homo sapi
3	1252	100.0	1745	9	AF218074	AF218074 Homo sapi
4	1252	100.0	1788	6	AR231198	AR231198 Sequence
5	1252	100.0	1788	6	AR307979	AR307979 Sequence
6	1252	100.0	2443	10	MUSTAK1	D76446 Mouse mRNA
7	1252	100.0	2656	6	AR231192	AR231192 Sequence
8	1252	100.0	2656	6	AR307973	AR307973 Sequence
9	1252	100.0	2757	9	BC017715	BC017715 Homo sapi
10	1252	100.0	2769	6	AX377912	AX377912 Sequence
11	1252	100.0	2769	9	AB009356	AB009356 Homo sapi
12	1252	100.0	2785	6	E38397	E38397 NF-kappa B
13	1252	100.0	2850	9	AB009357	AB009357 Homo sapi
14	1252	100.0	2866	6	E38398	E38398 NF-kappa B
15	1252	100.0	3107	10	BC006665	BC006665 Mus muscu
16	1221	97.5	2812	5	XU092030	U92030 Xenopus lae
17	1215	97.0	2745	5	BC049005	BC049005 Xenopus l
18	1133	30.5	135308	2	AC114407	AC114407 Mus muscu
19	700	55.9	3349	3	AF199466	AF199466 Drosophil
20	700	55.9	3386	3	AY051953	AY051953 Drosophil
21	599	47.8	32564	2	AC014558	AC014558 Drosophil
22	599	47.8	164942	3	AC011758	AC011758 Drosophil
23	599	47.8	302540	3	AE003571	AE003571 Drosophil
24	461.5	36.9	174275	5	BX119992	BX119992 Zebrafish
25	461.5	36.9	192888	2	BX005296	BX005296 Danio rer
26	392.5	31.3	3138	9	HSMSTM	Z48615 H.sapiens M
27	392.5	31.3	3454	6	AX337846	AX337846 Sequence
28	392.5	31.3	3454	9	HARNAMLK2	X90846 H.sapiens m
29	379	30.3	1269	6	AX653368	AX653368 Sequence
30	379	30.3	1269	8	AY156511	AY156511 Oryza sat
31	377	30.1	1130	6	AX653396	AX653396 Sequence
32	376	30.0	1445	5	AF265343	AF265343 Danio rer
33	375.5	30.0	8125	3	AF481923	AF481923 Dictyoste
34	374.5	29.9	8892	3	AY170918	AY170918 Dictyoste
35	374	29.9	1267	8	AY133876	AY133876 Arabidops
36	374	29.9	1798	8	AY046026	AY046026 Arabidops
37	373	29.8	1254	8	BT000587	BT000587 Arabidops
38	373	29.8	1631	8	AY085535	AY085535 Arabidops
39	373	29.8	1657	8	AY156510	AY156510 Oryza sat
40	373	29.8	1795	8	AY125513	AY125513 Arabidops
41	372.5	29.8	3931	9	AF251442	AF251442 Homo sapi
42	372	29.7	1429	10	AB049732	AB049732 Mus muscu
43	372	29.7	3079	10	BC023718	BC023718 Mus muscu
44	372	29.7	3146	10	AB049731	AB049731 Mus muscu
45	371.5	29.7	3092	6	AX675565	AX675565 Sequence

# ALIGNMENTS

RESULT 1

```

E38399      E38399      1704 bp      DNA      linear      PAT 31-JAN-2002
LOCUS      NP-kappa B activation inhibitory drug targeting TAK1 and method for
DEFINITION identifying the same.
ACCESSION  E38399
VERSION     E38399.1 GI:18626979
KEYWORDS    JP 2000197500-A/5.
SOURCE      unidentified
ORGANISM    unclassified.
REFERENCE   1 (bases 1 to 1704)
AUTHORS     Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.
TITLE       NP-kappa B activation inhibitory drug targeting TAK1 and method for
JOURNAL     Patent: JP 2000197500-A 5 18-JUL-2000;
COMMENT     OS Unidentified
            PN JP 2000197500-A/5
            PD 18-JUL-2000
            PE 04-FEB-1999 JP 1999026803
            PR NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI
            PI HASEGAWA
            PC C12Q1/48,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC
            PC C12Q1/02.
            PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09,
            PC C12R1:91).
            PC C12N5/00,C12N15/00,(C12N15/00,C12R1:91)
            CC Strandedness: Double;
            CC Topology: Linear;
            FH Key
            FT Location/Qualifiers
            FT 1..1704
            FT /organism="Unidentified".
FEATURES    source
            1..1704
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
BASE COUNT  511 a 381 c 401 g 411 t
ORIGIN
Alignment Scores:
Pred. No.:      1.3e-117      Length:      1704
Score:          1252.00      Matches:      228
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels: 0
DB:              6          Gaps: 0
US-09-830-144-2_COPY_76_303 (1-228) x E38399 (1-1704)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 226 GTAGAGCTCGGCAGTATCCCGTGTGAACCATCTAATATGTAAAGCTTTATGGAGCC 285
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 286 TGCITGATCAGTGTCGTGTGATGAATATCTGAAGGGGCTCTTTAATAAATGTG 345
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 346 CTGCATCGTGTGAACCATTTGCCATATTACTGTCTGCCAGCAATGAGTGTGTGTTA 405
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 406 CAGTGTTCACCAAGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 465
Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 466 GACCTGAACACCAACTTACTGCTGTTCCAGGGGGAGAGTCTTAAAAATTGTGAT 525
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120

```

```

526 TTTGGTACACGCTGTGACATTCAGACACACATGACCAATAACAGGGGAGTGTGCTTGG 585
121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
586 ATGGCACCTCAAGTTTTTGAAGGTAGTAATTACAGTGAATAATGTGACGCTCTTCAGCTGG 645
141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
646 GGTATTATTCTTTGGGAAGTGATAACGCGCGGAACCCCTTTGATGAGATTGGTGGCCCA 705
161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
706 GCITTCGATCATGTGGGCTGTTCATANTGGTACTCGACCACTGATAAAAAATTATA 765
181 ProLysProIleGlnSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
766 CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGTGCTTAAGATCCTCCACGCGCCT 825
201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
826 TCAATGAGGAATTTGAAAAATAATGACTCATCTGATGCGGTACTTTCCAGGAGCAGAT 885
221 GluProLeuGlnTyrProCysGln 228
886 GAGCCATTACAGTATCCTTGTCTAG 909
RESULT 2
AB009358      1705 bp      mRNA      linear      PRI 20-JAN-2001
LOCUS      Homo sapiens mRNA for TGF-beta activated kinase 1c, complete cds.
DEFINITION  AB009358
ACCESSION   AB009358
VERSION     AB009358.2 GI:8978251
KEYWORDS    tAK1c; TGF-beta activated kinase 1c.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS     Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE       TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
JOURNAL     NF-kappa B-inducing kinase-independent mechanism
MEDLINE     Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
PUBMED      98153801
PUBMED      9480845
REFERENCE   2 (sites)
AUTHORS     Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.
TITLE       Alternative splicing and gene structure of the transforming growth
JOURNAL     factor beta-activated kinase 1
MEDLINE     Biochim. Biophys. Acta 1517 (1), 46-52 (2000)
PUBMED      20568765
PUBMED      11118615
REFERENCE   3 (bases 1 to 1705)
AUTHORS     Sakurai,H., Shigemori,N., Hasegawa,K., Sugita,T. and Guesdon,F.
TITLE       Direct Submission
JOURNAL     Submitted (01-DEC-1997) Francois Guesdon, University of Sheffield,
            Royal Hallamshire Hospital, Division of Molecular and Genetic
            Medicine, Functional Genomics Group, Glossop road, Sheffield S10
            2UF, United Kingdom (E-mail:f.a.guesdon@sheffield.ac.uk,
            Fax:44-114-271-3846)
COMMENT     On Jul 8, 2000 this sequence version replaced gi:2924627.
FEATURES    Location/Qualifiers
            1..1705
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="pT7TAK1c"
            /cell_line="HeLa"
            1..1557
            /note="TAK1c"
            /codon_start=1
            /product="TGF-beta activated kinase 1c"
            /protein_id="BAA25027.2"
            /db_xref="GI:8978252"
CDS

```

/translation="MSTASAASSSSSSAGEMIEAPSOVLNFEIDYKEIIEVEVVG  
GAFGVCKAKWRKDVAKQIESSEKAFIVEQLRSVNHNPNIKLYGACLNVPCL  
VMEYAEGLSLYNVLHGEPYPTTAHMSWCLQCSQGVAYLHSMQPKALHRLDKPP  
NLLVAGGTVLKICDFGTACDIQTHMTNNKGAAMPAVEFGSNYSKCDVFSWGII  
LWEVITRKPPDEIGGPAFRIMVHNGTRPLKNLPKPIESLMTKWCSDPFSORPS  
MEEVITKMTLHMYFGADEPLQYPCOYSDRGNSATSTGFMIDIASINTSNKSDTN  
MEQVPATNDITIKRLESKLLKNOAQKQESGRLSLGASRGSSVESLPPTSEKRSADM  
SEIETARTATTAAYSKPRGHEKTASFGNILDVPEIVLSNGSQPRRSITDITVNGTEP  
QGVSSRSSPSVRMTTSGPTSEKPSHPWTPDDSTDTNGSDNSIPMAYLTLDHQLQ  
ARTSCRTGPG"

BASE COUNT 511 a 381 c 402 g 411 t

## ALIGNMENT SCORES:

Pred. No.: 1.3e-117 Length: 1705  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x AB009358 (1-1705)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
Db 226 GTAGAGCTTCGGCAGTATCCGCTGTGAACCATCTCTATATGTAAGCTTTATGAGCC 285  
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
Db 286 TCGTTGAATCCAGTGTCTTGTGTATGTAATATGCTGAAGGGGGCTCTTTATATAATGTG 345  
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
Db 346 CTGCATGTGTCTGAACCATTCATATATCTGCTGCCACGCAATGAGTGTGGTTTA 405  
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
Db 406 CAGTGTTCCTCAAGAGTGGCTTATCTTCACAGCATGCACCAACCAAGCGCTTAATTCACAGG 465  
Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
Db 466 GACCTGAAACCAACCAACTTACTGTGTGTGTCAGGGGGGACAGTCTTAAAAATTTGTGAT 525  
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120  
Db 526 TTTGTACAGCTGTGACATTCAGACACACATGACCAATACCAAGGGAGTGCTGCTTGG 585  
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
Db 586 ATGGCACCTGAGTTTGTGAAGTAGTAAATACAGTGAATAATGTGACGTCTTCAGCTGG 645  
Qy 141 GlyIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160  
Db 646 GGTATATTCTTTGGGAAGTGATAACGCGTCGGAAACCTTTGTATGAGATTGGTGGCCCA 705  
Qy 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
Db 706 GCTTTCGGAATCATGTGGGCTGTTTCAATAGTGACTCCGACCACTGATAAAAAATTTA 765  
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200  
Db 766 CCTAAGCCCATTGAGAGCTGATGACTGTGTGTGTGTCTTAAGATCCCTCCAGCGCCCT 825  
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
Db 826 TCAATGGAGGAAATGTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 885  
Qy 221 GluProLeuGlnTyrProCysGln 228  
Db 886 GAGCCATTACAGTATCCTTGTGCAG 909  
RESULT 3  
AF218074

LOCUS AF218074 1745 bp mRNA linear PRI 18-DEC-2000  
DEFINITION Homo sapiens TGF beta-activated kinase splice variant d (TAK1)  
mRNA, complete cds.  
ACCESSION AF218074  
VERSION AF218074.1 GI:6746614  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (Bases 1 to 1745)  
AUTHORS Dempsey, C.E., Sakurai, H., Sugita, T. and Guesdon, F.  
TITLE Alternative splicing and gene structure of the transforming growth  
factor beta-activated kinase 1  
JOURNAL Biochim. Biophys. Acta 1517 (1), 46-52 (2000)  
MEDLINE 20568765  
PUBMED 11118615  
REFERENCE 2 (Bases 1 to 1745)  
AUTHORS Dempsey, C.E. and Guesdon, F.  
TITLE Direct Submision  
JOURNAL Submitted (21-DEC-1999) Division of Molecular and Genetic Medicine,  
University of Sheffield, School of Medicine, Glossop Road,  
Sheffield S10 2UF, United Kingdom  
FEATURES  
source  
1. 1745  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
gene  
1. 1745  
/gene="TAK1"  
CDS  
7. 1482  
/gene="TAK1"  
/note="TAK1d"  
/codon\_start=1  
/product="TGF beta-activated kinase splice variant d"  
/protein\_id="AAF27652.1"  
/db\_xref="GI:6746615"  
/translation="MSTASAASSSSSSAGEMIEAPSOVLNFEIDYKEIIEVEVVG  
GAFGVCKAKWRKDVAKQIESSEKAFIVEQLRSVNHNPNIKLYGACLNVPCL  
VMEYAEGLSLYNVLHGEPYPTTAHMSWCLQCSQGVAYLHSMQPKALHRLDKPP  
NLLVAGGTVLKICDFGTACDIQTHMTNNKGAAMPAVEFGSNYSKCDVFSWGII  
LWEVITRKPPDEIGGPAFRIMVHNGTRPLKNLPKPIESLMTKWCSDPFSORPS  
MEEVITKMTLHMYFGADEPLQYPCOYSDRGNSATSTGFMIDIASINTSNKSDTN  
MEQVPATNDITIKRLESKLLKNOAQKQESGRLSLGASRGSSVESLPPTSEKRSADM  
SEIETARTATTAAYSKPRGHEKTASFGNILDVPEIVLSNGSQPRRSITDITVNGTEP  
QGVSSRSSPSVRMTTSGPTSEKPSHPWTPDDSTDTNGSDNSIPMAYLTLDHQLQ  
ARTSCRTGPG"  
BASE COUNT 526 a 375 c 412 g 432 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.33e-117 Length: 1745  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0  
US-09-830-144-2\_COPY\_76\_303 (1-228) x AF218074 (1-1745)  
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
Db 232 GTAGAGCTTCGGCAGTATCCGCTGTGAACCATCTCTATATGTAAGCTTTATGAGCC 291  
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
Db 292 TCGTTGAATCCAGTGTCTTGTGTATGTAATATGCTGAAGGGGGCTCTTTATATAATGTG 351  
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
Db 352 CTGCATGTGTGTAACCATTCGCTATATGCTGCTGCCACGCAATGAGTGTGGTTTA 411  
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80

```

Db      412 CAGTGTTCACAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 471
Qy      81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db      472 GACCTGAACACCAACTTACTGCTGGTTCAGGGGAGAGTCTTCTAAAAAATTTGTGAT 531
Qy      101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTIP 120
Db      532 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACAAAGGGAGGCTGCTGG 591
Qy      121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTIP 140
Db      592 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATACAGTGAATAATCTGACGCTTCAGCTGG 651
Qy      141 GlyIleLeuLeuTIPGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db      652 GGTATTATCTTTGGGAAGTGATTAACGCTGCGAAACCCCTTGTGATGAGATTGGTGGCCCA 711
Qy      161 AlaPheArgIleMetTIPAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db      712 GCCTTCGCAATCATGCGGCTGTTCATATGTTACTCGACACACATGATAAAAAATTTA 771
Qy      181 ProLysProIleGluSerLeuMetThrArgCysTIPSerLysAspProSerGlnArgPro 200
Db      772 CCTAAGCCCATTTGAGACCTGATGACTCGTTGTTGGTCTAAAGATCCTTCCAGCGCCT 831
Qy      201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db      832 TCAATGAGGAAATTTGAAATAATGACTCACTTGATGGGTACTTTCCAGGAGCAGAT 891
Qy      221 GluProLeuGlnTyrProCysGln 228
Db      892 GAGCCATTACAGTATCCTTTGTTCAG 915

RESULT 4
LOCUS      AR2311198              1788 bp      DNA      linear      PAT 20-DEC-2002
DEFINITION Sequence 14 from patent US 6451617.
ACCESSION  AR2311198
VERSION     AR2311198.1  GI:27272086
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 1788)
AUTHORS     Ono,K., Ohtomo,T. and Tsuchiya,M.
TITLE       Method of screening TGF-beta. inhibitory substances
JOURNAL     Patent: US 6451617-A 14 17-SEP-2002;
FEATURES    Location/Qualifiers
             source
             1..1788
             /organism="unknown"
BASE COUNT  551 a 394 c 413 g 430 t
ORIGIN
Alignment Scores:
Pred. No.:      1.37e-117      Length:      1788
Score:          1252.00      Matches:      228
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            6              Gaps:          0

US-09-830-144-2_COPY_76_303 (1-228) x AR2311198 (1-1788)

Qy      1 ValGluLeuArgGlnLeuSerArgValAlaHisProAsnIleValLysLeuTyrGlyAla 20
Db      232 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCCTAATATTGTAAAGCTTTATGGAGCC 291
Qy      21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
Db      292 TCGTTGAATCCAGTGTCTGTGATGGAATATCTGAAGGGGCTCTTTATATAATGTG 351
Qy      41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTIPCysLeu 60

```

```

Db      352 CTGATGGTCTGAACCACTTGCCATATTATCTGCTGCCACCGCAATGAGTGTGGTITTA 411
Qy      61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db      412 CAGTGTTCACAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 471
Qy      81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db      472 GACCTGAACACCAACTTACTGCTGGTTCAGGGGAGAGTCTTCTAAAAAATTTGTGAT 531
Qy      101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTIP 120
Db      532 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACAAAGGGAGGCTGCTGG 591
Qy      121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTIP 140
Db      592 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATACAGTGAATAATCTGACGCTTCAGCTGG 651
Qy      141 GlyIleLeuLeuTIPGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db      652 GGTATTATCTTTGGGAAGTGATTAACGCTGCGAAACCCCTTGTGATGAGATTGGTGGCCCA 711
Qy      161 AlaPheArgIleMetTIPAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db      712 GCCTTCGCAATCATGCGGCTGTTCATATGTTACTCGACACACATGATAAAAAATTTA 771
Qy      181 ProLysProIleGluSerLeuMetThrArgCysTIPSerLysAspProSerGlnArgPro 200
Db      772 CCTAAGCCCATTTGAGACCTGATGACTCGTTGTTGGTCTAAAGATCCTTCCAGCGCCT 831
Qy      201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db      832 TCAATGAGGAAATTTGAAATAATGACTCACTTGATGGGTACTTTCCAGGAGCAGAT 891
Qy      221 GluProLeuGlnTyrProCysGln 228
Db      892 GAGCCATTACAGTATCCTTTGTTCAG 915

RESULT 5
LOCUS      AR307979              1788 bp      DNA      linear      PAT 12-JUN-2003
DEFINITION Sequence 14 from patent US 6551840.
ACCESSION  AR307979
VERSION     AR307979.1  GI:31698736
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 1788)
AUTHORS     Ono,K., Ohtomo,T. and Tsuchiya,M.
TITLE       Method of screening TGF-beta.-inhibiting substances
JOURNAL     Patent: US 6551840-A 14 22-APR-2003;
FEATURES    Location/Qualifiers
             source
             1..1788
             /organism="unknown"
BASE COUNT  551 a 394 c 413 g 430 t
ORIGIN
Alignment Scores:
Pred. No.:      1.37e-117      Length:      1788
Score:          1252.00      Matches:      228
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:            6              Gaps:          0

US-09-830-144-2_COPY_76_303 (1-228) x AR307979 (1-1788)

Qy      1 ValGluLeuArgGlnLeuSerArgValAlaHisProAsnIleValLysLeuTyrGlyAla 20
Db      232 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCCTAATATTGTAAAGCTTTATGGAGCC 291

```

```

QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
Db 232 TGCCTTGAATCCAGTGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATATG 351
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
Db 352 CTGCATGTGTGTGAACCATTCGCAATATTATATCTGCTGCCACGCAATGAGTGTGTTA 411
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 412 CAGTGTTCCTCCAGGAGTGTCTTATCTTCACACATGCAACCCAAAGCGCTAATTCACAG 471
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 472 GACCTGAAACCAACCAACTTACTGTGTGTGAGGGGGACAGTCTCAAAAATTTGTGAT 531
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
Db 532 TTTGTGTACAGCGCTGTGACATTCAGACACACATGACCAATACCAAGGGGAGTCTGCTTG 591
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
Db 592 ATGGCACCTGAAGTTTTTGAAGGTAGTATTAATACAGTGAATAATGTGACGCTTTCAGTGG 651
QY 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 652 GGTATTATCTTTGGGAAGTGTATACGGGTGGAAACCTTTGATGAGATGTGTGGCCCA 711
QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 712 GCTTTCGGAATCATGTGGGTGCTTCATAATGTGTACTGCACCACTGATATAAATAATTA 771
QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
Db 772 CCTAAGCCCATGTAGAGCTGATGACTCGTGTGTGGTCTAAAGATCCTTCCAGCGCCCT 831
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 832 TCATGGAGGAATGTGAAATAATGATCTCACTGTATGCGGTACTTTTCCAGGAGCAGAT 891
QY 221 GluProLeuGlnTyrProCysGln 228
Db 892 GAGCCATTACAGTATCCTTGTCAG 915

RESULT 6
MUSTAK1
LOCUS Mouse mRNA for TAK1 (TGF-beta-activated kinase), complete cds.
DEFINITION D76446
ACCESSION D76446
VERSION 1
KEYWORDS TAK1; TGF-beta-activated kinase; protein kinase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Yanaguchi, K., Shirakabe, K., Shibuya, H., Irie, K., Oishi, I., Ueno, N.,
Taniguchi, I., Nishida, E. and Matsumoto, K.
Identification of a member of the MAPKK family as a potential
mediator of TGF-beta signal transduction
Science 270 (5244), 2008-2011 (1995)
96123277
8533096
2 (bases 1 to 2443)
Matsumoto, K.
Direct Submission
Submitted (18-OCT-1995) Kunihiro Matsumoto, Faculty of Science,
Nagoya University, Department of Molecular Biology; Furou-chou,
Chikusa-ku, Nagoya, Aichi 464-01, Japan
(E-mail: g44177aenucc.cc.nagoya-u.ac.jp, Tel: 052-789-3000,
Fax: 052-789-3001)
Location/Qualifiers
1. .2443
source

```

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
157. .1896
CDS
/codon_start=1
/product="TAK1 (TGF-beta-activated kinase)"
/protein_id="BAA11184.1"
/db_xref="GI:1167506"
/translation="MSTASAASSSSSSSSSEMEIAPSOVLNFEEDYKEIEVEEVVGR
GAFVVCCKAKRWAKDVAIKQIESERSEKAFIVELRQLSRVNHENIVKLYGACLNVPCL
VMEYAGGSLYNVLHGAEPLPYTAAHAMSWCLOCSGVAYLHSMOPKALHRLDKLPP
NLLVAGGTVLKI CDGFTACDIOHTMTNNKGAAMWAPEVEGNSYSEKCDVFSWGLI
LWEIVTRKPFDEIGGPAFRIMAVHNGTRPLIKNLPKPIESLMTRCWSKDPSPRS
MEELVIMTHLMRYFPGADEPLQPCYSDEGQNSATSTGFMDDIATNTSNKSDTN
MEQPARNDITIKRLESKLKNQAQOESGRLSGLSAGSSVESLPTTSEGRKSDAM
SEIARIVATAGNQPRRSIQDLTITGTEPGQVSSRSSSPSRVMTTSGTSEKPAR
SHPTPPDSDTNGSDNSIPWYLTLHDLOLPAPCPNSKESMAVFECHCKMAQRYMK
VOTRIALLQKQELVAELDQDQNTSLVQEHKKLLDENKSLSTYYQQCKQLE
VIRSQQKQKQKQTS"
BASE COUNT 669 a 568 c 647 g 559 t
ORIGIN
Alignment Scores:
Pred. No.: 1,94e-117 Length: 2443
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-830-144-2_COPY_76_303 (1-228) x MUSTAK1 (1-2443)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 382 GTGAGCTCCGCGAGTTGTCGCTGTGAACCATCTTAACATTTGTCAAGTTGTACGAGGC 441
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
Db 442 TGCCTGAATCCAGTATGCTTGTGATGGAATATGACAGGGGGGCTCATTTGATATATGTG 501
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
Db 502 CTGATGTGTGTGAACCATTCCTTACTACTGCTCATGCCATGAGCTGTGTTGTTA 561
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 562 CAGTGTTCCTCCAGGAGTGTCTTACTGTCACAGCTGACGCCAACGCTGATTCACAGG 621
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 622 GACCTCAAGCCTCCAAACTTCTGCTGTGTCAGAGGAGCAGTCTTAAATAATCTGCAT 681
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
Db 682 TTTGGTACAGCTGTGTGATCCAAACACACATGACCAATAATAAAGGGAGTGTGCTTGG 741
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
Db 742 ATGGCGCTGAAGTGTTTGAAGGTAGCAATTTACAGTGAAGTGTGATGCTTTCAGCTGG 801
QY 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 802 GGTATTATCTCTGGGAAGTGTATACACGCCGGAACCTTCGATGATGATCGGTGGCCCA 861
QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 862 GCTTTCAATCATGTGGCTGTGTTCAATATGGCACTGACCAACCACTGATCAAAAATTA 921
QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
Db 922 CCTAAGCCCATGTAGAGCTTGTGATGACAGCTGTGTGTTGTTCTAAGGACCCATCTCAGCGCCCT 981
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220

```

Db 982 TCAATGAGGAAATTGTGAAATAATGACTCACTTGATGGTACTTCCCGAGGCGGAT 1041  
Qy 221 GluProLeuGlnTyrProCysGln 228  
Db 1042 GAGCCATTACAGTATCCTTGTGAG 1065  
RESULT 7  
AR2311192 AR2311192 2656 bp DNA linear PAT 20-DEC-2002  
LOCUS Sequence 3 from patent US 6451617.  
ACCESSION AR2311192  
VERSION AR2311192.1 GI:27272080  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2656)  
AUTHORS Ono,K., Ohtomo,T. and Tsuchiya,M.  
TITLE Method of screening TGF-beta. inhibitory substances  
JOURNAL Patent: US 6451617-A 3 17-SEP-2002;  
FEATURES  
source  
1..2656  
/organism="unknown"  
BASE COUNT 778 a 557 c 620 g 701 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,13e-117 Length: 2656  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-830-144-2\_COPY\_76\_303 (1-228) x AR2311192 (1-2656)  
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
Db 408 GTAGAGCTTCGGCAGTATCCCGTGTGACCATCTTAATATTGTAAGCTTTATGGAGCC 467  
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40  
Db 468 TGCCTGAATCCAGGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATCTG 527  
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
Db 528 CTGCATGGTGTGAACCATTCGCCATATTACTCTGCCACGCAATGAGTTGGTGTTA 587  
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
Db 588 CAGTGTTCACAGAGTGGCTTATCTTTCACAGCATGCAACCCCAAGCGCTTAATTCACAGG 647  
Qy 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
Db 648 GACCTGAAACCCACCAAACTTACTGCTGGTTCAGGGGGACAGTTCTAAAAAATTTGTGAT 707  
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120  
Db 708 TTTGGTACAGCTGTGACATTCACACACATGACCAATAACAGGGGAGTGGCTGTGG 767  
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGlnLysCysAspValPheSerTyr 140  
Db 768 ATGGCAGCTGAGTTTTTGAAGTAGTAAATTCAGTGAAATATGAGCTTTCAGCTGG 827  
Qy 141 GlyIleLeuTyrGluValIleThrArgLysPropheAspGluIleGlyGlyPro 160  
Db 828 GGTATTATCTTTGGGAAGTATAACGGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 887  
Qy 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
Db 888 GCTTTCCGAATCATGTGGGCTGTTTCAATGATGTAACCGCTCGGACCACTGATATAAATTTA 947

Qy 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200  
Db 948 CCTAAGCCCATTCAGAGCCTGATGACTCGTTGTTGGTCTAAGATCTCTCCAGCGCCT 1007  
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
Db 1008 TCAATGAGGAAATTTGTGAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 1067  
Qy 221 GluProLeuGlnTyrProCysGln 228  
Db 1068 GAGCCATTACAGTATCCTTGTGAG 1091  
RESULT 8  
AR307973 AR307973 2656 bp DNA linear PAT 12-JUN-2003  
LOCUS Sequence 3 from patent US 6551840.  
ACCESSION AR307973  
VERSION AR307973.1 GI:31698730  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2656)  
AUTHORS Ono,K., Ohtomo,T. and Tsuchiya,M.  
TITLE Method of screening TGF-beta.-inhibiting substances  
JOURNAL Patent: US 6551840-A 3 22-APR-2003;  
FEATURES  
source  
1..2656  
/organism="unknown"  
BASE COUNT 778 a 557 c 620 g 701 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2,13e-117 Length: 2656  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-830-144-2\_COPY\_76\_303 (1-228) x AR307973 (1-2656)  
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
Db 408 GTAGAGCTTCGGCAGTATCCCGTGTGACCATCTTAATATTGTAAGCTTTATGGAGCC 467  
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40  
Db 468 TGCCTGAATCCAGGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATCTG 527  
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
Db 528 CTGCATGGTGTGAACCATTCGCCATATTACTCTGCCACGCAATGAGTTGGTGTTA 587  
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
Db 588 CAGTGTTCACAGAGTGGCTTATCTTTCACAGCATGCAACCCCAAGCGCTTAATTCACAGG 647  
Qy 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
Db 648 GACCTGAAACCCACCAAACTTACTGCTGGTTCAGGGGGACAGTTCTAAAAAATTTGTGAT 707  
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120  
Db 708 TTTGGTACAGCTGTGACATTCACACACATGACCAATAACAGGGGAGTGGCTGTGG 767  
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGlnLysCysAspValPheSerTyr 140  
Db 768 ATGGCAGCTGAGTTTTTGAAGTAGTAAATTCAGTGAAATATGAGCTTTCAGCTGG 827  
Qy 141 GlyIleLeuTyrGluValIleThrArgLysPropheAspGluIleGlyGlyPro 160  
Db 828 GGTATTATCTTTGGGAAGTATAACGGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 887  
Qy 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
Db 888 GCTTTCCGAATCATGTGGGCTGTTTCAATGATGTAACCGCTCGGACCACTGATATAAATTTA 947

161 AlapheArgilleMetTrrAlaValHisAsnGlyThrArgProProLeuLeuIleAsnLeu 180  
 Db GCTTTCCGAATCATGTGGCTGTTTCATATGATCTCACCACCACTGATAAAAAATTGA 947

181 ProLysProIleGluSerLeuMetThrArgCysTrrSerLysAspProSerGlnArgPro 200  
 Db CCTAAGCCCATTTGAGAGCTGATGACTCGTTGGTCTTAAGATCCCTTCCAGGCCCT 1007

201 SerMetGluGluIleValLysLeuMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db TCAATGAGGAAATGTGAAATAATGACTCATTGATCGGTACTTTCCAGGACAGAT 1067

221 GluProLeuGlnTyrProCysGln 228  
 Db GAGCCATTACAGTATCCTTTGTCAG 1091

RESULT 9  
 BC017715 2757 bp mRNA linear PRI 06-DEC-2001  
 LOCUS Homo sapiens, mitogen-activated protein kinase kinase 7,  
 DEFINITION clone MGC:21263 IMAGE:3906837, mRNA, complete cds.  
 ACCESSION BC017715  
 VERSION BC017715.1 GI:17389342  
 KEYWORDS MGC.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (Bases 1 to 2757)  
 AUTHORS Strausberg,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-DEC-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Prepared by: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAP Plate: 22 Row: 1 Column: 18  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 4507360.  
 Location/Qualifiers  
 1..2757  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="LocusID:6885"  
 /db\_xref="taxon:9606"  
 /clone="MGC:21263 IMAGE:3906837"  
 /tissue\_type="Uterus, leiomyosarcoma"  
 /clone\_lib="NIH\_MGC\_71"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 198..1937  
 /codon\_start=1  
 /product="mitogen-activated protein kinase kinase  
 7"  
 /protein\_id="RAH17715.1"  
 /db\_xref="GI:17389343"  
 /translation="MSTASAASSSSSSGAGEMIEAPSOVLNFEETDYKEIEVEEVVGR  
 GAFGVCKAKWKADVAIKQIESESERKAFIVELQSLRVNHPNIVKLYGLACLPVCL

CDS

BASE COUNT 819 a 566 c 647 g 725 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,22e-117 Length: 2757  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x BC017715 (1-2757)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
 Db GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTAATATTGTAAGCTTTATCGAGCC 482

QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
 Db TGCTTGATTCAGTGTGCTTTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 542

QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
 Db CTGATGTGTGTGAACCATTCCTATTTATCTGCTGCCACGCAATGAGTTGGTGTGTTA 602

QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
 Db CAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 662

QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAsp 100  
 Db GACCTGAAACCAACCAACTTACTGCTGTTGTCAGGGGGGACAGTTCTAAAAATTTGTGAT 722

QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120  
 Db TTGTGTACAGCTGTGACATTCAGACACATGACCAATACCAATACAGGGGAGTCTCTCTGG 782

QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
 Db ATGGCACCTGAAGTTTTTTGAAGTAGTAATTTACAGTGAATAATTTGACGTCTTCAGCTGG 842

QY 141 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160  
 Db GGTATTATTCTTTGGGAAGTATACGCGTCGGAACCTTTGATGAGATTGGTGGCCCA 902

QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleIleAsnLeu 180  
 Db GCTTTCCGAATCATGTGGCTGTTTCATATAATGGTACTCGACCACCACTGATAAAAAATTGA 962

QY 181 ProLysProIleGluSerLeuMetThrArgCysTrrSerLysAspProSerGlnArgPro 200  
 Db CCTAAGCCCATTTGAGAGCTGATGACTCGTTGGTCTTAAGATCCCTTCCAGGCCCT 1022

QY 201 SerMetGluGluIleValLysLeuMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db TCAATGAGGAAATGTGAAATAATGACTCATTGATCGGTACTTTCCAGGACAGAT 1082

QY 221 GluProLeuGlnTyrProCysGln 228  
 Db GAGCCATTACAGTATCCTTTGTCAG 1106

RESULT 10  
 AX377912  
 LOCUS AX377912 2769 bp DNA linear PAT 18-MAR-2002

```

DEFINITION Sequence 107 from Patent WO0212338.
ACCESSION AX377912
VERSION AX377912.1 GI:19573976
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Gillen,C., Wetzels,I., Wnendt,S., Weihe,E. and Schaefer,M.K.
TITLE Screening method
JOURNAL Patent: WO 0212338-A 107 14-FEB-2002;
Gruenenthal GmbH (DE)
FEATURES
    source
        1..2769
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
BASE COUNT 811 a 565 c 640 g 753 t
ORIGIN
Alignment Scores:
Pred. No.: 2,236-117 Length: 2769
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AX377912 (1-2769)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 388 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCCTAATATGTAAGCTTTATGGAGCC 447
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 448 TGCCTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTATATAATGTG 507
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 508 CTGCATGCTGCTGAACCATTTGCCATATATTACTGCTGCCACGCAATGAGTTGTTTGA 567
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 568 CAGTGTTCACCAAGAGTGGCTTATCTTCACAGATGCAACCAAGCGCTAATTCACAGG 627
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 628 GACCTGAACACCAACCACTTACTGCTGTTGCAGGGGGGACAGTTCTAATAAATTTGTGAT 687
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 688 TTTGGTACAGCTGTGACATTCACACACATACACATACCAATACCAAGGGAGTGTCTGTGG 747
QY 121 MetAlaProGluValPheGluGlySerAspTyrSerGlnLysCysAspValPheSerTrp 140
Db 748 ATGGCACTGAGTTTTTGAAGGTAGTAATTCAGTGAAAAATGTGACGCTTCAGCTGG 807
QY 141 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 808 GGTATTATTCTTTGGGAAGTGATAACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 867
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 868 GCTTTCGAATCATGTGGCTGTTCATAATGTGACTCGACCACTGATGATAAAAAATTTA 927
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 928 CCTAAGCCCATTTGAGACCTGATGACTGCTGTTGTTGTTTAAAGATCTTCCAGGCCCT 987
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220

```

```

Db 988 TCAATGAGGAAATTGTGAAAAATAATGACTCACTTGATGGGTACTTTCCAGGAGCAGAT 1047
QY 221 GluProLeuGlnTyrProCysGln 228
Db 1048 GAGCCATTACAGTATCCTTGTCTCAG 1071

RESULT 11
LOCUS AB009356 2769 bp mRNA linear PRI 04-MAR-1998
DEFINITION Homo sapiens mRNA for TGF-beta activated kinase 1a, complete cds.
ACCESSION AB009356
VERSION AB009356.1 GI:2924623
KEYWORDS TAK1a; TGF-beta activated kinase 1a.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
NF-kappa B-inducing kinase-independent mechanism
JOURNAL Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
MEDLINE 9480845
PUBMED 9480845
REFERENCE 2 (bases 1 to 2769)
AUTHORS Sakurai,H.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research
laboratory, Tanabe Seiyaku Co. Ltd.; 16-89, Kashima-3-chome,
Yodogawa-ku, Osaka, Osaka 532, Japan [E-mail:hsakurai@tanabe.co.jp,
Tel:++81 6 300 2571, Fax:++81 6 300 2593]
FEATURES
    source
        1..2769
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="pBstAK1a"
        /tissue="lung"
        /clone_lib="Lambda gt11"
        /notes="TAK1a"
        /codon_start=1
        /product="TGF-beta activated kinase 1a"
        /protein_id="BA225025.1"
        /db_xref="GI:2924624"
        /translation="MSTAAASSSSSSSSSAGEMTEAPSVLNFEEIDYKEIIEVEVGR
        GAFVYCKAKRWADVAIKQISESEKAFIVELRQLSRVNHPIVKLYGACLNVPCL
        VMEYAGGSLYNVLHGAELPYVTAHMSWCLQCGVAYLHSMOPKALIHRLDKPP
        NLLVAGGTVLKICDFGTACDIOTHNKNGSAAMWAPVEFGSNYSEKCDVRSWGI
        LMEVITRRPFDEIGGAPRIMAVNGTRPPLIKNLPKPIESIMTRCNKSKDPSQRES
        MEIVKIMHLMRYFFGADEPLQPCYDEGOSNATSTGSMDFDIASTNTRKSDTN
        MEQVPATNDTIKRLSKLKNQAKQSGESRGLSGASRGSSESLSLPTSGKMSADM
        SEIETARATNGNGRRRSIQDLTVTGPQVRSRSPSVRMITTSKPTSEKPTR
        SHPTPDDSDTDNGSDNSIPMAYLTLDHLOPLAPCPNSKESMAVTEQHCMAQEVYMK
        VOTEIALLLQRKQELVAELDQDEKQDQNTSRLVQEHKKLLDENKSLSTYYQQCKKQLE
        VIRSQOKRQGT5"
BASE COUNT 811 a 565 c 640 g 753 t
ORIGIN
Alignment Scores:
Pred. No.: 2,236-117 Length: 2769
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AB009356 (1-2769)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 388 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCCTAATATGTAAGCTTTATGGAGCC 447

```





LOCUS AB009357 2850 bp mRNA linear PRI 04-MAR-1998  
 DEFINITION Homo sapiens mRNA for TGF-beta activated kinase 1b, complete cds.  
 ACCESSION AB009357  
 VERSION AB009357.1 GI:2924625  
 KEYWORDS TAK1b; TGF-beta activated kinase 1b.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (sites)  
 AUTHORS Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.  
 TITLE NF-kappa B-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism  
 JOURNAL Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)  
 MEDLINE 98153801  
 PUBMED 9480845  
 REFERENCE 2 (bases 1 to 2850)  
 AUTHORS Sakurai,H.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research laboratory, Tanabe Seiyaku Co. Ltd.; 16-89, Kashima-3-chome, Yodogawa-ku, Osaka, Osaka 532, Japan (E-mail:hsakurai@tanabe.co.jp, Tel:++81 6 300 2571, Fax:++81 6 300 2593)  
 FEATURES  
 Location/Qualifiers  
 1..2850  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="pBStak1b"  
 /tissue type="lung"  
 /clone lib="Lambda gt11"  
 163..1983  
 /note="TAK1b"  
 /codon start=1  
 /product="TGF-beta activated kinase 1b"  
 /protein id="BAA25026.1"  
 /db\_xref="GI:2924626"  
 /translation="MSTASVSSSSSSAGEMTEAPSOVLNFEEDYKELEVERVGR  
 GAGVCKAKRWAKDAIKOLESESEKATIVELRQLSRVHNHVIKYLACLNPLVCL  
 VMEYAEGLSYNLVHGAEPPIYTAHAMSCLQSGVAILHSMQKALIHRLKPL  
 NLLVAGTGLKIDCFDAGDIQTHMTNKGSAWMAPEVFGNSYSEKCDVFSWGI  
 LWEVITRKPFDEIGGPAFRIMWAVHNGTRPPLIKLPKPIESIMTRCWSKQSPRS  
 MEEVITKMLHRYFGDADEPLQVCOYSDGQSNASATGSGFMDIATNTSNKSDTN  
 MEQVPANDTIKLKNQAKQOESGRSLSLGASRGSSVESLPTSSGRMSADN  
 SEIETARTATYAKPKRHKRTASFGNILDVPEIVISGNQPRRSIQDLTVTGEP  
 QVSSRSRSSVMITTSGTSEKPRSHPTWDDSDTNGSDNSIPMAKILDLHLQ  
 PLAPCPNSKESMAVFEQHCWAQVEYMKVTEIALLLQKQDELVAQLDQDEKQDNTSR  
 LVQEHKLLDENKSLSTYYQCKQLEVIRSQOQKQSGTS"  
 BASE COUNT 832 a 588 c 656 g 774 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,3e-117 Length: 2850  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatch: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-830-144-2\_COPY\_76\_303 (1-228) x AB009357 (1-2850)  
 QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyVala 20  
 Db 388 GTAGAGCTTGGCAGATTATCCCGTGTGAACATCCTTAATATTGTAAGCTTTATGGACC 447  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
 Db 448 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTCAAGGGGCTCTTTATATAATGTG 507  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
 Db 508 CTGCATGCTGTGAACCAATTGCCATATTACTGCTGCCACCGCAATGAGTGTGTGTTA 567

QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg 80  
 Db 568 CAGTGTTCCTCAAGAGTGGCTTATCTTCACAGATGCAACCAAGCGCTAATTCACAGG 627  
 QY 81 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
 Db 628 GACCTGAACACCAAACTTACTGCTGTTGCAGGGGGGACAGTCTTAAAAAATTGTGAT 687  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120  
 Db 688 TTGGTACACCCCTGTGACATTCACACACATGACCAATACAGGGGAGTGTGCTGG 747  
 QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140  
 Db 748 ATGGCACTGAAGTCTTTGAGGTAGTAATTACAGTGAATAATGTGACGCTTCAGCTGG 807  
 QY 141 GlyIleLeuLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160  
 Db 808 GGTATTATTCTTTGGGAAGTGATAACCGTCGGAACCCCTTGTGATGAGATTGGTGGCCA 867  
 QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
 Db 868 GCTTTCGAATCATGTGGCTGTTCATAATGTTACTCGACCACTGATATAAAATTTA 927  
 QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200  
 Db 928 CTTAAGCCCATTCAGAGCCTGATGACTGTTGTGTTAAAGATCCTTCCACGCGCCT 987  
 QY 201 SerMetGluGluLeuValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db 988 TCAATGAGGAAATTTGGAATAATGACTCACTTGATGCGGTACTTCCAGGAGCAGAT 1047  
 QY 221 GluProLeuGlnTyrProCysGln 228  
 Db 1048 GAGCCATTACAGTATCCTTGTGAG 1071  
 RESULT 14  
 LOCUS E38398 2866 bp DNA linear PAT 31-JAN-2002  
 DEFINITION NF-kappa B activation inhibitory drug targeting TAK1 and method for identifying the same.  
 ACCESSION E38398  
 VERSION E38398.1 GI:18626978  
 KEYWORDS JP 2000197500-A/4.  
 SOURCE unidentified  
 ORGANISM unclassified.  
 REFERENCE 1 (bases 1 to 2866)  
 AUTHORS Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.  
 TITLE NF-kappa B activation inhibitory drug targeting TAK1 and method for identifying the same.  
 JOURNAL Patent: JP 2000197500-A 4 18-JUL-2000;  
 COMMENT TANABE SEIYAKU CO LTD  
 OS Unidentified  
 PN JP 2000197500-A/4  
 PD 18-JUL-2000  
 PF 04-FEB-1999 JP 1999026803  
 PR  
 PI NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI HASEGAWA  
 PC C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC C12Q1/02,  
 PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09, PC C12R1:91),  
 PC C12N5/00,C12N15/00,(C12N15/00,C12R1:91)  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 FH Key Location/Qualifiers  
 FT source 1..2866  
 /organism="Unidentified".  
 Location/Qualifiers  
 1..2866  
 /organism="unidentified"

```

/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT      848 a      588 c      656 g      774 t
ORIGIN

Alignment Scores:
Pred. No.:      2,31e-117      Length:      2866
Score:          1252.00      Matches:      228
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              6      Gaps:      0

US-09-830-144-2_COPY_76_303 (1-228) x E38398 (1-2866)

Qy      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db      388 GTAGAGCTTCGGCAGTATCCGCTGATGACCATCTCTAATATTGTAAGCTTTATCGAGCC 447

Qy      21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db      448 TCGTTGAATCCAGTGTCTTGTGATGAATATGCTGAAGGGGGCTCTTTATATATATGTG 507

Qy      41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
Db      508 CTGCATGTGTGTAACCATTCATATTATATCTGCTGCCACGCAATGATGGTGTGTTA 567

Qy      61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db      568 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAAAGCGCTAATTCACAGG 627

Qy      81 AspLeuTyrProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db      628 GACCTGAACCAACCAACTTACTGCTGTTCAGGGGGACAGTCTTAAAAATTTGTGAT 687

Qy      101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
Db      688 TTTGGTACAGCTGTGACATTCAGACACACATGACCATACAGGGGAGTGTCTGTGG 747

Qy      121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
Db      748 ATGGCACTGAAAGTTTTTGAAGGTAGTAATTACAGTGAATAATGTACGCTTTCAGCTGG 807

Qy      141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyPro 160
Db      808 GGTATTATTCTTTGGGAAGTGAATACCGCTCGAAACCCCTTTGATGAGATTGGTGGCCCA 867

Qy      161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db      868 GCTTTCCGAATCATGTGGCTGTTTCATAATGTGTACTCGACCACTGATATAAAAAATTTA 927

Qy      181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
Db      928 CCTAAGCCCATGTAGAGCTGATGACTGTGTGTGGTCTAAAGATCCCTTCCAGCGCCCT 987

Qy      201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db      988 TCAATGAGGAAATGTGAAATAATGACTCACTGTATGCGGTACTTTCCAGGACGAGAT 1047

Qy      221 GluProLeuGlnTyrProCysGln 228
Db      1048 GAGCCATTACAGTATCTTGTGCAG 1071

RESULT 15
BC006665
LOCUS
DEFINITION
Mus musculus mitogen activated protein kinase kinase kinase 7, mRNA
(cDNA clone MGC:5989 IMAGE:3499247), complete cds.
ACCESSION
BC006665
VERSION
BC006665.1 GI:13879375
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

# REFERENCE

## AUTHORS

1 (bases 1 to 3107)

Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,

Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,

Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,

Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J.,

McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,

Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

Fahey J., Helton E., Kettman M., Madan A., Rodriguez S.,

Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,

Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,

Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

Butterfield Y.S., Krzywinski M.I., Skalski U., Smallos D.E.,

Schneerch A., Schein J.E., Jones S.J. and Marra M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

23389257

12477932

2 (bases 1 to 3107)

Direct Submission

Submitted (27-APR-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Jeffrey Green M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome

Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [amgbcn.tmc.edu](mailto:amgbcn.tmc.edu)

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 9 Row: h Column: 9

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis, GenomesCan gene prediction, Similarity but not identity

to protein.

Location/Qualifiers

1. 3107

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:5989 IMAGE:3499247"

/tissue\_type="Mammary tumor, C3(1)-Tag model. Infiltrating

ductal carcinoma. 5 month old virgin mouse."

/clone\_lib="NCI CGAP\_Mam6"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 3107

/gene="Map3k7"

/note="synonyms: Tak1, B430101B05"

/db\_xref="LocusID:26409"

/db\_xref="MGI:1346877"

148. .1968

CDS

FEATURES

source

gene

```
/codon_start=1
/product="Map3k7 protein"
/protein_id="AAH06665.1"
/db_xref="GI:13879376"
/db_xref="LocusID:26409"
/translation="MSTAGAASSSSSSASEMIEAPSOVLFNFEIDYKEIEVEVYGR
GARGVCKAKRWAKVAIKOIESERKAFIVELROLSRVNHENIVKLYGACLNPLVCL
VMEYAGGSLYNVLHGAELPYITRAHAMSNCLQCQGVAYLHSMQPKALLHRDLAPP
NLLLVAGGTVLKICDFGTACDIQTHMTNNKSAAWAPVFEGSNYSEKCDVFSWGI
MEVITRRKPFDRIGGPAFRIMWAVHNGTRPPLIKNLKPKIESLMTKMSKDPSPRS
MEVIMTHLMRYFPGADEPLQPCQYSDGQNSGATSGFMDIASNTSNKSDTN
MEQVPATNDITIKLESKLLKNAQKQESGRLLSIGASRGSSVESLPPTSSEKMSADM
SEIEARIVTAAYAKSKPKRHKRTASFGNILDVPEIVISNGQPRRSIQDLTVTGTPE
GOVSSRSSESVMITTSPTSEKPARSHPTWEDSDTSDNGSDNSIPMAYILDLHQIQ
PLAPCNSKESMAVFEQHKMAQYEMKVFETIAILLKOKELVIELDQDEKQOQNTSR
LVQEHKLLDENKSLSTYYQCKKQLEVIRSQQKRQGTs"
```

BASE COUNT 898 a 687 c 772 g 750 t

## ALIGNMENT SCORES:

Pred. No.:	2,53e-117	Length:	3107
Score:	1252.00	Matches:	228
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-830-144-2\_COPY\_76\_303 (1-228) x BC006665 (1-3107)

QY	1	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla	20
DB	373	GTGAGAGCTCGGCAGATTGTCGGTGTGAACCATCTTAACATTGTCAAGTTGTACGGAGCC	432
QY	21	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal	40
DB	433	TGCCTGAATCCAGTATGCTTGTGATGGAATATGCAGAGGGGGCTCATTTGATAATGTG	492
QY	41	LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu	60
DB	493	CTGCATGGTGTGAACCATTCCTTACTACACTGCTCATGCCATGAGCTGTGTGTTTA	552
QY	61	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg	80
DB	553	CAGTGTCCCAAGAGTGGCTTACTGACAGCATGACGCCCAAGCGCTGATTACAGG	612
QY	81	AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp	100
DB	613	GACCTCAAGCCTCCAAACTTGCTGCTGGTTCAGAGGGGACAGTTCTTAAATAATCTGGCAT	672
QY	101	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr	120
DB	673	TTTCGTACAGCTTGTGACATCCAAACACACATGACCAATAATAAGGGAGTGTGCTTGG	732
QY	121	MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr	140
DB	733	ATGGCCCTGAAGCTTTGAAGGTAGCAATTACAGTGAAAGTGTGATGCTTTCAGCTGG	792
QY	141	GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro	160
DB	793	GGTATTATCTCTGGGAAGTGATAACACGCCGGAACCCCTTCGATGAGATCGTGGCCCA	852
QY	161	AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu	180
DB	853	GCTTTCAAGATCATGTGGGCTGTTCTAATAATGGCACTGCACCACCATCATCAAAATTTA	912
QY	181	ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro	200
DB	913	CCTAAGCCCATTGAGAGCTTGATGACACGCTGTGGTCTAAGGACCCCATCTCAGCGCCT	972
QY	201	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp	220
DB	973	TCAATGGAGGAAATTTGTGAAATAATAATGACTCACTTGATGCGGTACTTCCCGAGCGGAT	1032

QY 221 GluProLeuGlnTyrProCysGln 228  
DB 1033 GAGCCATTACAGTATCTTGTCTAG 1056

Search completed: December 4, 2003, 11:39:56  
Job time : 4917.16 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 09:34:12 ; Search time 397.459 Seconds  
(without alignments)  
1548.515 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303

Perfect score: 1252  
Sequence: 1 VELRQLSRVNHPIVILYGA.....MTHLMRYFPGADEPLQLPQC 228

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US090310144/runat 03122003 122509 23313/app query.fasta\_1.654  
-DB=N\_Geneseq 19Jun03 -QPMF=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US090310144 @CIGN 1.0 @runat 03122003 122509 23313 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq 19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	100.0	1704	20	AAIX99698 Human TGF-beta act
2	1252	100.0	1788	20	AAIX56285 Human TAK1-6xHis e
3	1252	100.0	1959	18	AAI85095 Mouse transforming
4	1252	100.0	2443	18	AAI85094 Human TAK1 encodin
5	1252	100.0	2656	20	AAIX56279 Human TAK-1 nucleo
6	1252	100.0	2656	21	AAAX39105 Pain regulated cDN
7	1252	100.0	2769	24	ABL88437 Human TGF-beta act
8	1252	100.0	2785	20	AAIX99696 Human TGF-beta act
9	1252	100.0	2866	20	AAIX99697 Drosophila melanog
10	700	55.9	3367	23	ABL02489 Drosophila melanog
11	599	47.8	10997	23	ABL02488 Drosophila melanog
12	414	33.1	759	23	ABL08337 Pancreas cancer re
13	392.5	31.3	3454	24	ABL70018 Arabidopsis thalia
14	374	29.9	1224	21	AAIC3254 Arabidopsis thalia
15	373	29.8	1631	21	AAIC39537 Human NOV7, a mixe
16	371.5	29.7	3092	24	ABN85385 Human colon cancer
17	371	29.6	1063	22	AAH34976 cDNA encoding huma
18	371	29.6	1365	21	AAZ99727 cDNA encoding a hum
19	371	29.6	1706	21	AAI75674 cDNA encoding huma
20	371	29.6	2120	21	AAZ99726 Human protein enco
21	371	29.6	2191	22	AAH99263 Cardiovascular sys
22	371	29.6	2194	21	AAZ99734 Human survival reg
23	371	29.6	2220	21	AAZ93783 Cardiovascular sys
24	371	29.6	2254	21	AAZ99735 Novel protein kina
25	371	29.6	2403	22	AAI44701 Motif zipper conta
26	371	29.6	2592	24	ABQ60980 Human TGF-beta rec
27	371	29.6	2622	22	AAI75336 Human cervical can
28	371	29.6	3967	22	AAH73366 Cardiovascular sys
29	370	29.6	2069	21	AAZ99737 Cardiovascular sys
30	370	29.6	2272	21	AAZ99736 Drosophila melanog
31	363.5	29.0	3072	23	ABL04365 Human kinase (PKIN
32	361.5	28.9	3141	22	AAD18824 Human PKIN-12 cDNA
33	361.5	28.9	3538	24	AAI34309 Novel human gene.
34	355.5	28.4	3066	24	ABQ86165 cDNA encoding huma
35	355	28.4	2157	22	AAH46913 Novel human protei
36	355	28.4	3111	24	ABN86357 Nucleotide sequenc
37	355	28.4	3111	25	ABZ23269 Human kinase and p
38	355	28.4	3111	25	AAI49450 Novel human protei
39	355	28.4	3518	24	ABN86358 Polynucleotide seq
40	353.5	28.2	2732	25	ABZ23196 Polynucleotide seq
41	353.5	28.2	2830	25	ABZ23195 Human cDNA differe
42	353.5	28.2	3365	24	ABK84203 Polynucleotide seq
43	353.5	28.2	3365	25	ABZ23194 Human leucine zipp
44	353.5	28.2	3389	16	AAI01031 Human leucine-zipp
45	353.5	28.2	3389	18	AAI89349

## ALIGNMENTS

RESULT 1  
AAIX99698  
ID AAIX99698 standard; cDNA to mRNA; 1704 BP.  
XX  
AC AAIX99698;  
XX  
DT 18-OCT-1999 (first entry)  
XX  
DE Human TGF-beta activated kinase (TAK) 1c encoding nucleotide sequence.  
XX  
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;  
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;  
KW intracellular disease; atrophic dermatitis; psoriasis; viral infection;  
KW endotoxin shock; septicemia; human; hTAK1c; ss.  
XX  
OS Homo sapiens.  
XX

```

FH Key Location/Qualifiers
FT CDS 1..1704
FT /*tag= a
FT /product= "hTAK1c"
XX WO9940202-A1.
XX 12-AUG-1999.
XX 02-FEB-1999; 99WO-JP00422.
XX 30-OCT-1998; 98JP-0309316.
XX 06-FEB-1998; 98JP-0026003.
XX (TANA ) TANABE SEIYAKU CO.
XX
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX WPI; 1999-494298/41.
XX P-PSDB; AAY28998.
XX
XX Nuclear factor kappa B activation inhibitors, useful as preventives
XX for, e.g. autoimmune diseases
XX
XX Examples; Page 43-46; 49pp; Japanese.
XX
XX The invention provides a method for identifying or screening a nuclear
XX factor kappa B (NF-kB) activation inhibitor by examining the effect of a
XX test substance on modulating the function(s) of TGF-beta activated kinase
XX 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
XX treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
XX intractable diseases with inflammation (such as atrophic dermatitis and
XX psoriasis), viral infection, endotoxin shock, septicemia and others. The
XX present sequence represents the nucleotide sequence of human TAK1c
XX (hTAK1c) protein.
XX
XX Sequence 1704 BP; 511 A; 381 C; 401 G; 411 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 3.75e-133 Length: 1704
XX Score: 1252.00 Matches: 228
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 20 Gaps: 0
XX
XX US-09-830-144-2_COPY_76_303 (1-228) x AAX99698 (1-1704)
XX
XX 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
XX 226 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCCCTTAATATGTAAGCTTTATGAGCC 285
XX
XX 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
XX 286 TGGTTCGAATCCAGTGTCTGTGTGATGAATATGCTGAAGGGGGCTTTATATAATGTG 345
XX
XX 41 LeuHisGlyValGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
XX 346 CTGCATGGTGCTGAACCATTCGCATATATATCTGCTGCACCAATGAGTTGGTGTGTTA 405
XX
XX 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArg 80
XX 406 CAGTGTTCACAGGAGTGGCTTATCTTCACAGCATGCACCAACCGCTAATTCACAGG 465
XX
XX 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
XX 466 GACCTGAACCAACCAACTTACTGCTGGTGTGAGGGGGGACAGTCTTAAAAAATTTGTGAT 525
XX
XX 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
XX 526 TTGGTACAGCTGTGACATTCAGACACATGACCAATACCAAGGGGAGTGTGCTTGG 585
XX
XX 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
XX
XX
XX Db 586 ATGCACCTGAAGTTTTTGAAGGTAGTAATTACAGTAAAAAATGTGACGTTTCAGCTGG 645
XX
XX Qy 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
XX
XX Db 646 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTGATGAGATTGGTGGCCCA 705
XX
XX Qy 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
XX
XX Db 706 GCTTTCCGAATCATGTGGCTGTTCATAATGTGACTCGACCCACCATCATATAAAAAATT 765
XX
XX Qy 181 ProllysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
XX
XX Db 766 CCTAAGCCCATTCAGAGCCCTGATGCTGTGTGGTCTAAAGATCTTCCACGGCCCT 825
XX
XX Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
XX
XX Db 826 TCAATGGAGGAATTTGAAAAATAATGACTCATTGATCGGTACTTTCAGGAGCAGAT 885
XX
XX Qy 221 GluProLeuGlnTyrProCysGln 228
XX
XX Db 886 GAGCCATTACAGTATCCTTGTCTAG 909
XX
XX RESULT 2
XX AAX56285
XX ID AAX56285 standard; DNA; 1788 BP.
XX AC AAX56285;
XX XX 21-JUL-1999 (first entry)
XX DE Human TAK1-6xHis encoding DNA.
XX KW Human; TAK1; screening; inhibition; TGF-beta;
XX transforming growth factor beta; ss.
XX OS Homo sapiens.
XX OS Synthetic.
XX Key Location/Qualifiers
XX CDS 7..1779
XX /*tag= a
XX
XX WO9921010-A1.
XX 29-APR-1999.
XX 22-OCT-1998; 98WO-JP04796.
XX 22-OCT-1997; 97JP-0290188.
XX (CHUS ) CHUGAI SEIYAKU KK.
XX Ohtomo T, Ono K, Tsuchiya M;
XX WPI; 1999-312645/26.
XX P-PSDB; AAY09547.
XX
XX Screening for TGF- beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX
XX Example 1; Page 167-171; 195pp; Japanese.
XX
XX A method has been developed for screening for substances which inhibit
XX the binding of TAK1 polypeptide to TAK1 polypeptide. The method
XX comprises: (a) contacting the polypeptide in the presence of a sample;
XX and (b) detecting the amount of bound polypeptide, in which the sample
XX can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming
XX growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX indications e.g. as TGF-beta signal transduction inhibitors or
XX activators, or extracellular matrix protein production enhancement
XX inhibitors or activators, or cell proliferation prevention inhibitors or
XX activators, or monocyte migration inhibitors or activators, or

```

CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence encodes TAK1-6XHis from an example of  
 CC the present invention.

XX  
 SQ Sequence 1788 BP; 551 A; 394 C; 413 G; 430 T; 0 other;

## Alignment Scores:

Pred. No.: 4.01e-133 Length: 1788  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x AAX56285 (1-1788)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
 DB GTAGAGCTTCGGCAGTATCCCGGTGAACCATCTTAATATGTAAAGCTTTATGGAGCC 291  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
 DB TGCCTTGATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 351  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
 DB CTGCATGGTGTGAACCATTCCTATATATATCTGCCACGCAATGAGTTGGTGTTA 411  
 QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
 DB CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTACACAG 471  
 QY 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
 DB GACCTGAACCCACCAACTTACTGCTGGTTCAGGGGGGACAGTCTTAATAATTTGTGAT 531  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120  
 DB TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTGG 591  
 QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
 DB ATGGCACCCTGAAGTCTTTTGAAGGTAGTAATACAGTGAATAATGTGACGCTCTTCAGCTGG 651  
 QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160  
 DB GGTATTAATCTTTGGGAAGTGTATACCGCTCGAAACCCCTTTGATGAGATTTGGTGCCCA 711  
 QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
 DB GCTTTCCGAATCATGTGGCTGTTCAATAGTACTGACCAACCATCTGATAAAAAATTA 771  
 QY 181 ProLysProIleGlnSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200  
 DB CCTAAGCCCATGTGAGAGCCTGATGACTCGTGTGTGGTCTAAGATCCCTCCAGCGCCCT 831  
 QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 DB TCATATGAGGAATATGTGAATAATATGACTCATCTGTATGCGGTACTTTCCAGGAGCAGAT 891  
 QY 221 GluProLeuGlnTyrProCysGln 228  
 DB GAGCATTACAGTATCTCTTGTTCAG 915

## RESULT 3

AAAT85095  
 ID AAAT85095 standard; cDNA; 1959 BP.  
 XX  
 AC AAAT85095;  
 XX

19-NOV-1997 (first entry)  
 Human transforming growth factor-beta activated kinase TAK-1 cDNA.  
 TGF-beta; signal transduction; TGF-beta activated kinase;  
 MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;  
 protein kinase; ss.

Homo sapiens.

Key Location/Qualifiers  
 CDS 183..1922  
 /\*tag= a  
 /product= TAK-1

JP09163990-A.

24-JUN-1997.

27-SEP-1996; 96JP-0256747.

24-JUL-1996; 96US-0685625.

29-SEP-1995; 95JP-0253549.

(CHUS) CHUGAI PHARM CO LTD.

(UENO) UENO N.

WPI; 1997-380171/35.

P-PSDB; AAW27093.

DNA encoding transforming growth factor-beta-activated kinase, TAK-1  
 - useful for studying the TGF-beta signal transduction system

Claim 9; Page 13-15; 20pp; Japanese.

The present sequence encodes human transforming growth factor-beta  
 (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the  
 TAK-1 protein which is involved in the TGF-beta family signal  
 transduction system. TAK-1, also known as activator of MAPK Kinase  
 (AMK-1), is an enzyme which is activated by TGF-beta and bone  
 morphogenetic protein (BMP) and activates MAPK kinase by  
 phosphorylation.

Sequence 1959 BP; 571 A; 447 C; 481 G; 460 T; 0 other;

## Alignment Scores:

Pred. No.: 4.57e-133 Length: 1959  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 18 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x AAAT85095 (1-1959)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
 DB GTAGAGCTTCGGCAGTATCCCGGTGAACCATCTTAATATGTAAAGCTTTATGGAGCC 467  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
 DB TGCCTTGATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 527  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
 DB CTGCATGGTGTGAACCATTCCTATATATCTGCCACGCAATGAGTTGGTGTTA 587  
 QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
 DB CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTACACAG 647  
 QY 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100

Db 648 GACCTGAACACCAAACTTACTGCTGGTTGCGAGGGGAGCAGTCTTAAATAATTTGTGAT 707  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIleGlySerAlaAlaTyr 120  
 Db 708 TTTGGGTACAGCCGTGTGACATTCAGACACACATGACCAATAACAGGGAGTGTCTGTGG 767  
 QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140  
 Db 768 ATGGCACTGAAGTTTTTGAAGGTAGTAATTCAGTGAATAATGTGACGTCTTCAGCTGG 827  
 QY 141 GlyIleLeuLeuTyrGluValIleThrArgArgIysProPheAspGluIleGlyGlyPro 160  
 Db 828 GGTATTATTCTTTGGGAAGTGATTAACGCGTCGGAACCCCTTTCATGAGATTGTGGCCCA 887  
 QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
 Db 888 GCTTTCGAATCATGTGGGCTGTTTCATAATGGTACTGACCCACCATGATAAAAAATTTA 947  
 QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200  
 Db 948 CCTAAGCCCATTCAGACCCCTGATGACTCGTGTGTGTCTTAAGATCCTTCCAGGCGCCT 1007  
 QY 201 SerMetGluGluLeuValIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db 1008 TCAATGGAGGAAATTTGAAAAATAATGACTCACTTGATGGGTACTTTCCAGGAGCAGAT 1067  
 QY 221 GluProLeuGlnTyrProCysGln 228  
 Db 1068 GAGCCATTACAGTATCCTTGTGAG 1091

## RESULT 4

AAT85094  
 ID AAT85094 standard; cDNA; 2443 BP.

AC AAT85094;

19-NOV-1997 (first entry)

Mouse transforming growth factor-beta activated kinase TAK-1 cDNA.

TGF-beta; signal transduction; TGF-beta activated kinase;

MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;

protein kinase; ss.

Mus musculus.

Key Location/Qualifiers  
 CDS 157..1896  
 FT /\*tag= a  
 FT /product= TAK-1

JP09163990-A.

24-JUN-1997.

27-SEP-1996; 96JP-0256747.

24-JUL-1996; 96US-0685625.

23-SEP-1995; 95JP-0253549.

(CHUS) CHUGAI PHARM CO LTD.

(UENO) UENO N.

WEI; 1997-380171/35.

P-PSDB; AAW27092.

DNA encoding transforming growth factor-beta-activated kinase, TAK-1

- useful for studying the TGF-beta signal transduction system

Claim 2; Page 10-12; 20pp; Japanese.

The present sequence encodes mouse transforming growth factor-beta

(TGF-beta) activated kinase, TAK-1. The DNA is used to produce the

CC TAK-1 protein which is involved in the TGF-beta family signal  
 CC transmission system. TAK-1, also known as activator of MAPK Kinase  
 CC (AMK-1), is an enzyme which is activated by TGF-beta and bone  
 CC morphogenetic protein (BMP) and activates MAPK kinase by  
 CC phosphorylation.

XX  
 SQ Sequence 2443 BP; 669 A; 567 C; 647 G; 559 T; 1 other;

## Alignment Scores:

Pred. No.: 6,26e-133 Length: 2443  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservations: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 18 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x AAT85094 (1-2443)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
 Db 382 GTGGAGCTCCGGCAGTGTGCGGTGTGAACCATCTTAACATTGTCAAGTTGTACGGAGCC 441  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
 Db 442 TGCCTGAATCCAGTATGTCTTGTGATGGAATATGACAGGGGGGCTCATTTGATAATGTG 501  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
 Db 502 CTGCATGTGTCTGAACCATTTGCCCTTACTACACTGCTGCTCATGCCATGAGCTGTGTGTTA 561  
 QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
 Db 562 CAGTGTTCACAGAGGTGGCTTACCTGCACAGCATGCAGCCCAAGCGCTGATTCACAGG 621  
 QY 81 AspLeuLysProLysLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
 Db 622 GACCTCAAGCTCCAAACTTGTCTGTTGTCAGGAGGAGCAGTTCTAAAAATCTGGCAT 681  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120  
 Db 682 TTTGGTACAGCTTGTGACATCCAAACACACATGACCAATTAATAAGGGAGTGTCTGTGG 741  
 QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140  
 Db 742 ATGGCGCTCAAGTGTGTGAAGTAGCAATTACAGTGAAGTGTGATGCTTCTTCAGCTGG 801  
 QY 141 GlyIleLeuLeuTyrGluValIleThrArgArgIysProPheAspGluIleGlyGlyPro 160  
 Db 802 GGTATTATCTCTGGGAAGTGATAACACGCGCGAAACCCCTTCGATGAGATCGGTGGCCCA 861  
 QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
 Db 862 GCTTTCAGAAATCATGTGGGCTGTTCATATATGGCATCGCACCATCTGATCAAAATTTA 921  
 QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200  
 Db 922 CCTAAGCCCATTCAGAGCTTCATGACACGCTGTGGTCTTAAGGACCATCTCAGCGCCCT 981  
 QY 201 SerMetGluGluLeuValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db 982 TCAATGGAGGAAATTTGAAAAATAATGACTCACTTGATGGGTACTTCCAGGAGCGGAT 1041  
 QY 221 GluProLeuGlnTyrProCysGln 228  
 Db 1042 GAGCCATTACAGTATCCTTGTGAG 1065

## RESULT 5

AA556279

ID AA556279 standard; DNA; 2656 BP.

XX AC AA556279;

XX DT 21-JUL-1999 (first entry)



XX DE Human TAK1 encoding DNA.  
 XX KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
 XX KW transforming growth factor beta; ss.  
 XX OS Homo sapiens.  
 XX PH Key Location/Qualifiers  
 XX FT CDS 183..1922  
 XX FT /\*tag= a  
 XX PN WO9921010-A1.  
 XX PD 29-APR-1999.  
 XX PF 22-OCT-1998; 98WO-JP04796.  
 XX PR 22-OCT-1997; 97JP-0290188.  
 XX PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX PI Ohtomo T, Ono K, Tsuchiya M;  
 XX DR WPI; 1999-312645/26.  
 XX DR P-PSDB; AAY09542.  
 XX PT Screening for TGF- beta inhibitory substances, which are useful as  
 XX PT drugs for treatment of diseases relating to its disorder  
 XX PS Example 1; Page 150-154; 195pp; Japanese.  
 XX CC A method has been developed for screening for substances which inhibit  
 XX CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method  
 XX CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 XX CC and (b) detecting the amount of bound polypeptide, in which the sample  
 XX CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming  
 XX CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
 XX CC indications e.g. as TGF-beta signal transmission inhibitors or  
 XX CC activators, or extracellular matrix protein production enhancement  
 XX CC inhibitors or activators, or cell proliferation prevention inhibitors or  
 XX CC activators, or monocyte migration inhibitors or activators, or  
 XX CC physiological activity induction inhibitors or activators, or  
 XX CC immunosuppression inhibitors or activators, or amyloid beta protein  
 XX CC precipitation inhibitors or activators, and such substances can also be  
 XX CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 XX CC activity. The present sequence encodes human TAK1.  
 XX SQ Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 7.05e-133 Length: 2656  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0  
 US-09-830-144-2\_COPY\_76\_303 (1-228) x AAX56279 (1-2656)  
 QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTy-GlyAla 20  
 Db 408 GFAGAGCTTCGGCAGTATCCCGTGTGAACCAATCTAATATTGTAAAGCTTTATGAGGCC 467  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
 Db 468 TGTCTGATCCAGTGTCTTGTGATGGATATGCTGAAGGGGCTCTTATATATATGTG 527  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
 Db 528 CTGCATGTGTGTAACCAATTCGCATATTATATATATATATATATATATATATATATAT 587  
 QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80

Db 598 CAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCCGCTAATTCACAGG 647  
 QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
 Db 648 GACCTGAAACCAACCAACTTACTGCTGGTTGCAGGGGGACAGTTCTAAAAAATTTGTGAT 707  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIleTrp 120  
 Db 708 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAGGGGAGTGCTGCTTGG 767  
 QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
 Db 768 ATGGCACCCTGAAGTTTGAAGGTAGTAATTCACAGTGAATAATGACAGCTCTTCAGCTGG 827  
 QY 141 GlyIleLeuLeuTrpGluValIleThrArgArgLysPropheAspGluIleGlyGlyPro 160  
 Db 828 GGTATTATTCTTTGGGAAGTGATAACGGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 887  
 QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
 Db 888 GCTTTCCGAATCATGTGGGCTGTTCAATAATGGTACTCGACCACCACTGATAAAAATTTTA 947  
 QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200  
 Db 948 CCTAAGCCCATTTGAGAGCTGATGACTGCTGTGGTCTAAGATCCTTCCAGGCCCT 1007  
 QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db 1008 TCAATGGAGGAAATTTGTGAAAAATAATGACTCACTTGATGCGGTACTTTCCAGGACAGAT 1067  
 QY 221 GluProLeuGlnTyrProCysGln 228  
 Db 1068 GAGCCATTACAGTATCCTTGTGAG 1091  
 RESULT 6  
 ID AAA39105 standard; DNA; 2656 BP.  
 AC AAA39105;  
 DT 04-SEP-2000 (first entry)  
 DE Human TAK-1 nucleotide sequence SEQ ID NO:1.  
 KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;  
 KW screening; signal transduction; inhibition; inflammatory cytokine;  
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;  
 KW antiinflammatory; suppression; ds.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 CDS 183..1922  
 FT /\*tag= a  
 FT /product= "TAK-1"  
 PN WO200023610-A1.  
 PD 27-APR-2000.  
 PF 21-OCT-1999; 99WO-JP05817.  
 PR 21-OCT-1998; 98JP-0299962.  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;  
 DR WPI; 2000-339707/29.  
 DR P-PSDB; AAY91000.  
 PT Method for screening inhibitors of TAK1 signal transduction for

PT suppression of inflammatory cytokine production and use as  
 PT antiinflammatory agents -

PS Example 1; Page 73-80; 100pp; Japanese.

XX The present invention describes a method for screening compounds for  
 CC inhibition of inflammatory cytokine signal transduction by contacting  
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition  
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds  
 CC for inhibition of inflammatory cytokine signal transduction in which the  
 CC inhibition of TAK1 phosphorylation is selected for; and drug  
 CC compositions for the treatment of inflammatory disorders containing as  
 CC active component an inflammatory cytokine signal transduction inhibitor.  
 CC TAK1 is an essential component of the signalling process which results  
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),  
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used  
 CC for the selection of effective antiinflammatory agents. The present  
 CC sequence encodes human TAK-1, which is used in the exemplification of the  
 CC present invention.

SQ Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;

# Alignment Scores:

Pred. No.: 7.05e-133 Length: 2656  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x AAA39105 (1-2656)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20  
 Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACATCCTAATATTTGAAAGCTTTATGGAGCC 467  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
 Db 468 TGCTTGATCCAGTGTCTGTGTGATGGAATATGCTGAAGGGGCTTTTATATATATG 527  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
 Db 528 CTGCATGGTCTGAACCATTCGCCATATATATCTCTCCACGCAATGAGTTGGTGT 587  
 QY 61 GluCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80  
 Db 588 CAGGTCTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAAATTCACAGG 647  
 QY 81 AspleuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
 Db 648 GACTGNAACACCAACTTACTGTCTGTGTGAGGGGGGACAGTCTTAAAAATTTGAT 707  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLeuGlySerAlaAlaTrp 120  
 Db 708 TTTGGTACAGCGTGTGACATTCACACACATGACCAATACCAAGGGGAGTGTCTGTG 767  
 QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluIysCysAspValPheSerTrp 140  
 Db 768 ATGGCACTGAAGTTTGTGAAGGTAGTAAATACAGTGAATAATGTACGTTCTTCAGCTGG 827  
 QY 141 GlyIleIleLeuTrpGluValIleThrArgArgIysProPheAspGluIleGlyGlyPro 160  
 Db 828 GGTATTATTCTTTGGGAAGTGATACCGTTCGGAACCTTTGATGAGATTGGTGGCCCA 887  
 QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
 Db 888 GCTTTCGAATCATGTGGCTGTTCATAATGTACTCGACCACTGATAAAAAATT 947  
 QY 181 ProIysProIleGluSerLeuMetThrArcCysTrpSerLysAspProSerGlnAtrp 200  
 Db 948 CCAAGCCCAATGAGACCGTATGACATCGTTGTGGTAAAGATCCTTCCAGCGCCCT 1007  
 QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220

Db 1008 TCAATGGAGGAATTTGAAATAATGACTCACTTGATCGGTACTTTCAGGAGCAGAT 1067

QY 221 GluProLeuGlnTyrProCysGln 228

Db 1068 GAGCATTACAGTATCCTTGTTCAG 1091

## RESULT 7

ABL88437  
 ID ABL88437 standard; cDNA; 2769 BP.  
 XX ABL88437;  
 XX AC ABL88437;  
 XX 16-MAY-2002 (first entry)  
 DT Pain regulated cDNA sequence 80.  
 DE Pain; analgesic; gene therapy; neurological disorder;  
 KW neurodegenerative disease; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX WO200212338-A2.  
 PN 14-FEB-2002.  
 PD 03-AUG-2001; 2001WO-EP09011.  
 PF 03-AUG-2000; 2000DE-1037759.  
 PR  
 XX (CHEF ) GRUENENTHAL GMBH.  
 XX  
 XX Gillen C, Wetzel S, Wenden S, Weihe E, Schaefer MK;  
 PI WPI; 2002-257469/30.  
 XX P-PSDB; ABB85033.  
 DR  
 PT Identifying pain-regulating compounds, useful for treating chronic pain  
 PT and for diagnosis, by measuring binding of compounds to specific  
 PT peptides and proteins  
 XX  
 XX Claim 1; Fig 44; 213pp; German.  
 XX The invention relates to identifying pain-regulating substances (A)  
 CC comprises (i) incubating a test substance with a cell (or preparation  
 CC from it) that has synthesised a peptide or protein (B) and (ii) measuring  
 CC either binding of the test substance to (B) or some functional parameter  
 CC that is altered by this binding. The method is useful for identifying  
 CC pain-regulating substances (A) with analgesic activity. (A) along with  
 CC nucleic acid (ABL88411-ABL88441) that encode proteins (B),  
 CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the  
 CC nucleic acid; antibodies against (B); cells that express (B) and agents  
 CC that bind to (B), are all useful for treating pain, particularly chronic  
 CC pain, including use in gene therapy. The same materials can also be used  
 CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The  
 CC present sequence is that of a polynucleotide of the invention.  
 XX Sequence 2769 BP; 811 A; 565 C; 640 G; 753 T; 0 other;

## Alignment Scores:

Pred. No.: 7.48e-133 Length: 2769  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 24 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x ABL88437 (1-2769)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20  
 Db 388 GTAGAGCTTCGGCAGTTATCCCGTGTGAACATCCTAATATTTGAAAGCTTTATGGAGCC 447

```

QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
DB 448 TGCTTGAATCCAGTGGCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATG 507
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
DB 508 CTGCATGGTGGTGAACCAATGCCATATTACTGCTGCCACGCAATGAGTTGGTCTTTA 567
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg 80
DB 568 CAGTGTTCGAAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTTAATTCACAG 627
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 628 GACCTGAACCAACCAACCTTACTGGTTGCGAGGGGGACAGTTCTTAAAAATTTGTGAT 687
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
DB 688 TTGGGTACAGCCTGTGCATTCAGACACACATGACCAATAACAAGGGAGTGGCTTGG 747
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
DB 748 ATGGCACCTGAAGCTTTTGAAGGTAGTAATTACAGTGAATAAATGTGACGCTTCAGCTGG 807
QY 141 GlyIleIleLeuTrpGluValIleThrArgGlyProPheAspGluIleGlyGlyPro 160
DB 808 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTGATGAGTTGGTGGCCCA 867
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeu 180
DB 868 GCTTTCCGAATCATGTGGGCTGTTCATAATGTTACTGACCAACCTGATGAATAAATTTA 927
QY 181 ProLysProLysGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 928 CCTAAGCCCATTCAGAGCCTGATGACTCGTGTGTGGTCTAAGATCCTCCAGCGCCCT 987
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
DB 988 TCAATGAGGAATTTGGAATAATGACTCACTTGTATGCGGTACTTTCCAGAGCAGAT 1047
QY 221 GluProLeuGlnTyrProCysGln 228
DB 1048 GAGCCATTACAGTATCCTTGTCTAG 1071

RESULT 8
AAAX99696
ID AAAX99696 standard; cDNA to mRNA; 2785 BP.
AC
XX AAAX99696;
XX
DT 18-OCT-1999 (first entry)
DE Human TGF-beta activated kinase (TAK) 1a encoding nucleotide sequence.
XX
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1a; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 163..1902
FT /*tag= a
FT /product= "hTAK1a"
XX
EN WO9940202-A1.
XX
XX 12-AUG-1999.
XX
XX 02-FEB-1999; 99WO-JP00422.
XX
XX 30-OCT-1998; 98JP-0309316.
PR

```

```

PR 06-FEB-1998; 98JP-0026003.
XX (TANA ) TANABE SEIYAKU CO.
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX WPI; 1999-494298/41.
XX P-PSDB; AAY28996.
XX Nuclear factor kappa B activation inhibitors, useful as preventives
PT for, e.g. autoimmune diseases
XX
XX Examples; Page 35-39; 49pp; Japanese.
XX
XX The invention provides a method for identifying or screening a nuclear
XX factor kappa B (NF-kB) activation inhibitor by examining the effect of a
XX test substance on modulating the function(s) of TGF-beta activated kinase
XX 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
XX treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
XX intractable diseases with inflammation (such as atrophic dermatitis and
XX psoriasis), viral infection, endotoxin shock, septicemia and others. The
XX present sequence represents the nucleotide sequence of human TAK1a
XX (hTAK1a) protein.
XX
XX Sequence 2785 BP; 827 A; 565 C; 640 G; 753 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 7,54e-133 Length: 2785
XX Score: 1252.00 Matches: 228
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 20 Gaps: 0
XX
XX US-09-830-144-2_COPY_76_303 (1-228) x AAX99696 (1-2785)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyVala 20
DB 388 GTAGAGCTTCGGCAGTATTCCTGCTGACCATCTTAATATTGTAAGCTTTATGGAGCC 447
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
DB 448 TGCTTGAATCCAGTGGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATG 507
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
DB 508 CTGCATGGTGGTGAACCAATTCATATATCTCAGCATGCAACCAAGCGCTTAATTCACAG 567
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg 80
DB 568 CAGTGTTCGAAGGAGTGGCTTATCTTCAGCATGCAACCAAGCGCTTAATTCACAG 627
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 628 GACCTGAACCAACCAACCTTACTGCTGGTTCAGGGGGACAGTTCTTAAAAATTTGTGAT 687
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
DB 688 TTTGGTACAGCCTGTGACATTCAGACACATGACCAATAACAAGGGAGTGGCTTGG 747
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
DB 748 ATGGCACCTGAAGCTTTTGAAGGTAGTAATTACAGTGAATAAATGTGACGCTTCAGCTGG 807
QY 141 GlyIleIleLeuTrpGluValIleThrArgGlyProPheAspGluIleGlyGlyPro 160
DB 808 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTGATGAGTTGGTGGCCCA 867
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeu 180
DB 868 GCTTTCCGAATCATGTGGGCTGTTCATAATGTTACTGACCAACCTGATGAATAAATTTA 927
QY 181 ProLysProLysGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200

```

Db 928 CCTAAGCCATTGAGAGCCTGATGACTCGTTGTTGTTCTAAAGATCCTTCCAGAGCCCT 987  
 QY 201 SerMetGluGluIleValIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db 988 TCAATGAGGAAATTTGAAATTAATGACTCACTGATGGGTACTTCCAGGAGCAGAT 1047  
 QY 221 GluProLeuGlnTyrProCysGln 228  
 Db 1048 GAGCCATTACAGTATCCTTGTCTAG 1071

## RESULT 9

AAAX99697  
 ID AAX99697 standard; cdna to mRNA; 2866 BP.

XX  
 AC AAX99697;

XX  
 DT 18-OCT-1999 (first entry)

XX Human TGF-beta activated kinase (TAK) 1b encoding nucleotide sequence.

XX Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;  
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;  
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;  
 KW endotoxin shock; septicemia; human; hTAK1b; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FH 163..1983  
 FT CDS  
 FT /\*tag= a  
 FT /product= "hTAK1b"

XX W09940202-A1.

XX 12-AUG-1999.

XX 02-FEB-1999; 99WC-JP00422.

XX 30-OCT-1998; 98JP-0309316.

XX 06-FEB-1998; 98JP-0026003.

XX (TANA ) TANABE SEIYAKU CO.

XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;

XX WPI; 1999-494298/41.

XX P-PSDB; AAY28997.

XX Nuclear factor kappa B activation inhibitors, useful as preventives  
 for, e.g. autoimmune diseases

XX Examples; Page 39-43; 49pp; Japanese.

XX The invention provides a method for identifying or screening a nuclear  
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a  
 CC test substance on modulating the function(s) of TGF-beta activated kinase  
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
 CC intractable diseases with inflammation (such as atrophic dermatitis and  
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The  
 CC present sequence represents the nucleotide sequence of human TAK1b  
 CC (hTAK1b) protein.

SQ Sequence 2866 BP; 848 A; 588 C; 656 G; 774 T; 0 other;

## Alignment Scores:

Pred. No.: 7.86e-133 Length: 2866  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x AAX99697 (1-2866)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20  
 Db 388 GTAGACCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGAAAGCTTTATGAGCC 447  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyCysSerLeuTyrAsnVal 40  
 Db 448 TGCCTTGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTC 507  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
 Db 508 CTGCATGGTCTGAACCATTCATTAATTAATCTGCTGCCACCAATGATGTTGGTGTTA 567  
 QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIlyslAlaLeuIleHisArg 80  
 Db 568 CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 627  
 QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyThrValLeuIlyslCysAsp 100  
 Db 628 GACCTGAAACCCCAAACTTACTGCTGTTGCAGGGGGGACAGTTCTTAAAAATTTGTGAT 687  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120  
 Db 688 TTTGGTACAGCCTGTGACATTCAGACACATGACCAATACCAAGGGGAGTGTCTGTGG 747  
 QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140  
 Db 748 ATGGCACCTGAAGTTTTCAGAGTAGTAATTACAGTGAAAAATGTGACGCTCTTCAGCTGG 807  
 QY 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160  
 Db 808 GGTATTATTCTTTGGGAAGTGATAACGGCTCGGAAACCCCTTTGATGAGATTGTTGCCCA 867  
 QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProIleIleLysAsnLeu 180  
 Db 868 GCTTTCCGNAATCATGTGGGCTGTTTATATGTTGTTCTGACCCACTGATAAAAAATTTA 927  
 QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200  
 Db 928 CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTTGTTCTTAAAGATCCTTCCAGGCGCCT 987  
 QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db 988 TCAATGAGGAGAAATTTGAAATAATGACTCACTTGTATCGGTACTTTCCAGGAGCAGAT 1047  
 QY 221 GluProLeuGlnTyrProCysGln 228  
 Db 1048 GAGCCATTACAGTATCCTTGTCTAG 1071

## RESULT 10

ABL02489

ID ABL02489 standard; cdna; 3367 BP.

XX ABL02489;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1949.

XX Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-09231.

XX 23-MAR-2000; 2000US-191637P.

```
PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE ) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX XX
DR WPI; 2001-656860/75.
XX P-PSDB; ABB58386.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX PS Claim 1; SEQ ID NO 1949; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3367 BP; 938 A; 898 C; 856 G; 675 T; 0 other;

Alignment Scores:
Pred. No.: 1,69e-69 Length: 3367
Score: 700.00 Matches: 130
Percent Similarity: 73.01% Conservative: 35
Best Local Similarity: 57.52% Mismatches: 57
Query Match: 55.91% Indels: 4
DB: 23 Gaps: 3

US-09-830-144-2_COPY_76_303 (1-228) x ABL02489 (1-3367)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly----- 19
Db 1112 GAGTGAAGCAGTTGTGGCGGTGAGACCCCGAACAATCATCTGTCACGGGATATCC 1171
QY 20 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 1172 TCGTACCAGCAGCCACCTACCTGATAATGGAGTTCGCCGAGGTGGATCGTGCACAAC 1231
QY 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 59
Db 1232 TTCCTTTCACGGC---AAGGTGAAGCGCGCATATCTCTGCGCCACGCGCATGCTGGCG 1288
QY 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 1289 CGCAATGTCAGAGGTCTGGCATATTGTCATGCCATGACGCGCAACCACTAATACAT 1348
QY 80 ArgAspLeuIleProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
Db 1349 CGCGACGTGAAGCGCTGTAACCTGCTTGCACCAACAGGAGCGCAATCTGAAGATATGC 1408
QY 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySerAlaAla 119
Db 1409 GACTTCGCGACGCGTGGCGGACAACTGACCATGATGACCAACAATCGCGGAGTGGCGCT 1468
QY 120 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 139
Db 1469 TGGATGCGCGCCGAGGCTTCGAGGGTCCAGTATACGAGAGTGTGACATTTTAGC 1528
QY 140 TrpGlyIleIleLeuTrpGluValIleThrArgLysPropheAspGluIleGlyGly 159
Db 1529 TGGGCCATGTTCTATGGAGGTTCTGTCCAGGAAGCAGCCCTTTAAAGGCATCGCAAT 1588
QY 160 ProAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsn 179
Db 1799 TTTTTCCTTTCACGGC---AAGGTGAAGCGCGCATATCTCTGCGCCACGCGCATGCTGGCG 1848
```

```
Db 1589 ---GCCTACCATCCAGTCCAGTGAAGATCTACAAGGTGAACGCCCGCTGTCTGACCAT 1645
QY 180 LeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArg 199
Db 1646 TGCCCAAGCGCATCGAGGACCTGATGACCGCTCTGTGAAACCGTGCCTGAGGATCGC 1705
QY 200 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAla 219
Db 1706 CCGTGCATGCAGTACATAGTGGCGCTTATGCACGAGATCGTCAAGGACTATACGGGGCG 1765
QY 220 AspGluProLeuGlnTyr 225
Db 1766 GACAAGCGCTGGAATAC 1783

RESULT 11
ABL02488
ID ABL02488 standard; cDNA; 10997 BP.
XX
AC ABL02488;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 1946.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
P-PSDB; ABB58385.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 1946; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 10997 BP; 3237 A; 2362 C; 2382 G; 3016 T; 0 other;

Alignment Scores:
Pred. No.: 3,38e-57 Length: 10997
Score: 599.00 Matches: 130
Percent Similarity: 50.61% Conservative: 35
Best Local Similarity: 39.88% Mismatches: 57
Query Match: 47.84% Indels: 105
DB: 23 Gaps: 5
```

US-09-830-144-2\_COPY\_76\_303 (1-228) x ABL02488 (1-10997)

```
QY      2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGly----- 19
Db      2652 GAGGTGAAGAGTTGTCGGCGGTGAAGCACCACATATCGCTCTGCACGGGATATCC 2711
QY      20 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db      2712 TCGTACCAGCAGCCACCTACCTGATATGAGTTCGCCGAGGTGATCGTGCACAC 2771
QY      40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 59
Db      2772 TTCCITTCACGGC---AAGGTGAAGCCCGCATATCTCTGCCCCACCCATGACTGGCG 2828
QY      60 LeuGlnCysSerGln----- 64
Db      2829 CGCCATGTGCAGAGTAGTGTCTGTCAGGGGGGGGGGTTATAACGTAGGGTCAA 2888
QY      64 ----- 64
Db      2889 CCCCGTAGCGCTTCCATCCAAATTCGTGAAATATCTGTGAAATATGACTCAATGG 2948
QY      64 ----- 64
Db      2949 ATTGTGGTAATTCACATGCTTGCATGATTACTTAAACAATTTCAATAGCTATTCGCAT 3008
QY      64 ----- 64
Db      3009 AACCAGAAATTAATAATAGTAATTTGTTAAAGTTAGCCCTAATTACATTTCCATTG 3068
QY      65 ----- GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisAr 80
Db      3069 TTTTCTCTCTAGGCTGGCATATTTGTCATCCATGACGCCCAAAACCACTAATCATCG 3128
QY      80 gAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAs 100
Db      3129 CCAGGTGAAGCCGCTGAACCTCTCTTGACCAACAGGAGCACCAATCTCAAGATATGCA 3188
QY      100 pPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTr 120
Db      3189 CTTCCGACGGTGGCGGCAAGTCGACCATGATGACCAATCCGCGGAGTGGCGCTTG 3248
QY      120 pMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTr 140
Db      3249 GATGGCGCCGAGGCTTCGAGGCTCCAAAGTATACGGAGAAGTGTGACATTTTAGCTG 3308
QY      140 pGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyLysPr 160
Db      3309 GGCCATGTCTATGGAGGTTCTGTCCAGGAAGCAGCCCTTTAAAGGCATCGACAAT-- 3366
QY      160 oAlaPheArgIleMetTyrAlaValHisAsnGlyThr----- 172
Db      3367 -GCCTACACCATCCAGTGAAGATCTACAAGGTGC-GTCCCTCCAAATCACTCTTTTCC 3424
QY      173 ----- ArgProProLeuIleLysAsnLe 180
Db      3425 ATCGAGCTTATCGAGATCTGCTCTCGAGGTGAACGCCCGCGCTGCTGACCACTTG 3484
QY      180 uProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPr 200
Db      3485 CCCCAAGCGCATCGAGGACCTGATACCGCTGTGGAAGAAAGCGTGCAGGATCGGCC 3544
QY      200 oSerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGluValAs 220
Db      3545 GTCCATGAGTACATAGTGGCGTTATGCACGAGATCGTCAAGGACTATACGGGGCGGA 3604
QY      220 pGluProLeuGlnTyr 225
Db      3605 CAAGGCCCTGGAATAC 3620
```

RESULT 12

ABL08337

. ID ABL08337 standard; cDNA; 759 BP.

```
XX      ABL08337;
XX      26-MAR-2002 (first entry)
XX      Drosophila melanogaster expressed polynucleotide SEQ ID NO 19493.
XX      Drosophila; developmental biology; cell signalling; insecticide;
XX      pharmaceutical; gene; ss.
XX      Drosophila melanogaster.
XX      WO200171042-A2.
XX      27-SEP-2001.
XX      23-MAR-2001; 2001WO-US09231.
XX      23-MAR-2000; 2000US-191637P.
XX      11-JUL-2000; 2000US-0614150.
XX      (PEKE ) PE CORP NY.
XX      Venter JC, Adams M, Li PWD, Myers BW;
XX      WPI: 2001-656860/75.
XX      P-PSDB; ABB64234.
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX      Claim 1; SEQ ID NO 19493; 21pp + Sequence Listing; English.
XX      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention is
XX      useful in developmental biology and in elucidating cell signalling and
XX      cell-cell interactions in higher eukaryotes for the development of
XX      insecticides, therapeutics and pharmaceutical drugs. The invention
XX      discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX      sequences (ABL01840-ABL16175) and the encoded proteins
XX      (ABB57737-ABB72072).
XX      The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pct_sequences.
XX      SQ Sequence 759 BP; 226 A; 162 C; 198 G; 173 T; 0 other;

Alignment Scores:
Pred. No.: 1,17e-37 Length: 759
Score: 414.00 Matches: 82
Percent Similarity: 59.69% Conservative: 35
Best Local Similarity: 41.84% Mismatches: 53
Query Match: 33.07% Indels: 26
DB: 23 Gaps: 4

US-09-830-144-2_COPY_76_303 (1-228) x ABL08337 (1-759)
QY      2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCys 21
Db      160 GAGATCTACCAAGTGCACAAAGGCCACCATGTCACATAGTTGAGCTCTACGGCACATCG 219
QY      22 LeuAsnProValCys-----LeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db      220 AGGCACGAGGATCGCCCTCTGTGATGAATTCGTAGACGGTGAATCTCTGTCCAGT 279
QY      40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 59
Db      280 TTCTGTGCAC---GCGAAAGCAAGCAAGTATTTCGATGCCACCCCTTCACTGGCG 336
QY      60 LeuGlnCysSerGlnGlyValAlaTyrIleHisSerMetGlnProLysAlaLeuIleHis 79
Db      337 CATCAGATCGCTCAGGGCATAGCCTATCTGCATGTCATGCGATGCGAGCCGAAAGCAGTAATTCAT 396
```

```

QY      80 ArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
Db      397 CGGATATAAGCCACTCAATACACTGCTATGCGAGAGGACTCAAACTGAAGATTGC 456
QY      100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
Db      457 GATTTCGGAACGTGTTGTGGACCTATCCCAATCGATATCGTATCGCAATGCGGCACCTGCAGA 516
QY      120 TrpMetAlaProGlu----- 124
Db      517 TACAAAGCGCCGAGGTAAGGAGTGTGTTGATTTCAAATCCAAATCAATGAATAATCAAC 576
QY      125 -----ValPheGluGlySerAsnTyrSerGluLysCysAspVal 137
Db      577 CAACCAACCGCGCTTCAAAAGTTTCTACAAGAAATAAACCCGATGAAGAAAGTGGCATGTG 636
QY      138 PheSerTrpGlyIleLeuLysValIleThrArgLysProPheAspGluIle 157
Db      637 TATAGTTGGCTATTACTTTTGGAAATATGTCGCGAAGGAGCCATTGAGCAATAT 696
QY      158 GlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArg 173
Db      697 ---AATACGCTTTTGAACGTATACATGCTATTATTAATGAAGCAAGAGA 741

RESULT 13
ABL70018
ID ABL70018 standard; DNA; 3454 BP.
XX
AC ABL70018;
XX
DT 15-MAY-2002 (first entry)
XX
DE Pancreas cancer related gene sequence SEQ ID NO:8355.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.

```

```

PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237117P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
(aval-) AVALON PHARM.
YOUNG PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;
WPI, 2002-188264/24.
Screening for anti-neoplastic agent involves exposing cells to a
chemical agent to be tested for anti-neoplastic activity, and
determining a change in expression of a gene of a signature gene set -
Claim 1; SEQ ID 8355; 44pp; English.
The present invention describes a method (M1) for screening for an
anti-neoplastic agent. The method involves exposing cells to a chemical
agent to be tested for anti-neoplastic activity, determining a change in
expression of at least one gene (I) of a signature gene set, where (I)
comprises a sequence (S) selected from 8447 sequences (given in ABL61664
to ABL70110), or is at least 95% identical to (S), where a change in
expression is indicative of anti-neoplastic activity. (I) has cytostatic
activity and can be used in gene therapy. M1 can be used for screening
an anti-neoplastic agent, and can be used for producing a product which
is the data collected with respect to the anti-neoplastic agent as a
result of M1, and the data is sufficient to convey the chemical
structure and/or properties of the agent. M1 can be used in the
treatment of cancer such as colon, breast, stomach, lung, thyroid,
oesophageal, ovarian, kidney, prostate or pancreatic cancer,
adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
carcinoma, papillary carcinoma and Wilm's tumour.
SQ Sequence 3454 BP; 594 A; 1217 C; 1136 G; 507 T; 0 other;
Alignment Scores:
Pred. No.: 2.93e-34 Length: 3454
Score: 392.50 Matches: 88
Percent Similarity: 56.68% Conservative: 35
Best Local Similarity: 40.55% Mismatches: 77
Query Match: 31.35% Indels: 17
DB: 24 Gaps: 6
US-09-830-144-2_COPY_76_303 (1-228) x ABL70018 (1-3454)
QY      2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db      721 GAAGCCCGGCTCTTTGGAGCCCTGCGACACCCCAACATAATTGCCCTTAGGGGCGCCGTC 780
QY      22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db      781 CTCAACCCCCCACACCTCTGCTAGTAGAGTAGATGATGATGATGATGATGATGATGATGATG 840
QY      40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
Db      841 GTGCTGGCAGGTGCGCGGGTGCACCTCACGTG-----CTGTCAACTGGGCT 888

```

```
QY 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProTyrAlaLeuHis 79
Db 889 GTTCAGGTGGCGGGGCGATGAACCTACCAATGATGCCCTGTGCCATCATCCAC 948
QY 80 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92
Db 949 CGGACCTCAAGTCCATCAATCATCTGATCTGGAGGCCATCGAGAACCAACCTCGCA 1008
QY 93 GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThrHisMet 111
Db 1009 GACACGGTGTCTAAGATACAGGACTTCGGCTCCCGCGAGTGGCAACAGACCAAC 1068
QY 112 ThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyr 131
Db 1069 ATGACGCTCGGGGACCTACGCTGATGGCGCGGAGTTATCCGCTCTCCCTCTTC 1128
QY 132 SerGluLysCysAspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArg 151
Db 1129 TCCAAAGACGATGATGCTCGAGCTTCGGGTGCTGCTGGGAGCTGCTGACGGGGGAG 1188
QY 152 LysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaVal---HisAsn 170
Db 1189 GTCCCTACCGTGAGATC-----GAGCCCTTGGCGTGGCGTATGGCGTGTATGAAT 1242
QY 171 GlyThrArgProProLeuLysAsnLeuProLysProLysProLysLeuMetThrArg 190
Db 1243 AAGCTGACGTGCCCATTCCTCCACGTGCGCGAGCCCTTGGCGGCTCTCTGGAGAA 1302
QY 191 CysTrpSerLysAspProSerGlnArgProSerMetGluLysValLys 207
Db 1303 TGCTGGACCCAGACCCCGGCGCCAGATTTGCGTAGCATCTTGAAG 1353

RESULT 14
AAC43254
ID AAC43254 standard; DNR; 1224 BP.
XX AC AAC43254;
XX DT
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 38594.
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132883.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
```



```

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145132.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147439.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 8.8e-33
Score: 374.00
Percent Similarity: 57.80%
Best Local Similarity: 39.45%
Query Match: 29.87%
DB: 21
Gaps: 9

US-09-830-144-2_COPY_76_303 (1-228) x AAC43254 (1-1224)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 523 GAGGTGCTATGCTTGTCTAATTTGAAGCACCACCAACATTGTGAGGTTTCATTTGGTGCATGT 582
Qy 22 LeuAsnProVal-----CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 583 CGCAGCCAAATGGTGTGTATAGTACTGATACGCCAAAGAGGTTTCACTGAGGCAG 642
Qy 40 ValLeu-----HisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAla 55
Db 643 TTTTGTAGTAGGACAGACAGACCGGCGTCTCTTTGAAGTTA-----GCT 687
Qy 56 MetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLys 75
Db 688 GTTAAACAGCGTTTGGATGTTGTAGGGTATGCTTATGTCATGGA-----CGC 738
Qy 76 AlaLeuIleHisArgAspLeuLysProAsnLeuLeuValAlaGlyThrVal 95
Db 739 AACTTCATACAGAGATCTCAAGTCAGATACCTTCTCATCTCAGCAGATAGTCCATC 798
Qy 96 LeuLysIleCysAspPheGlyThrAla---CysAspIleGlnThrHis---MetThrAsn 113
Db 799 ---AAGATTGCAGATTTTGGTTGTCAGAAATTGAAGTTCAACCGAAGGAATGACACCA 855
Qy 114 AsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu 133
Db 856 GAAACTGGAACCTTACAGATGGATGGCTCCAGAGATGATACAGCATAGAGCCGTACAAATCAA 915
Qy 134 LysCysAspValPheSerTrpGlyIleIleTrpGluValIleThrArgArgLysPro 153
Db 916 AAGTGGATGTATAGTTTCGGGATTTGCTCGGGGTGAGTGCATATGACTCGATGT 975
Qy 154 PheAspGluIleGlyGly-----ProAlaPheArgIleMetTrpAlaValHisAsnGly 171
Db 976 TTCAGAACATGACAGCTGTACAGCGAGCGTTTCCGGTT-----GTAAACAGAGGA 1026
Qy 172 ThrArgProProGluIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCys 191
Db 1027 GTGGTCCAAACAGTCCCAACAGATGTCTCCGGTGTGAGTGCATATGACTCGATGT 1086
Qy 192 TrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db 1087 TGGATGCTAATCCAGAAAGTCCGTCATGTTTGTGGAGGTTGTGAAGCTGCT 1140

RESULT 15
AAC39537
ID AAC39537 standard; DNA; 1631 BP.

```

[illegible]

```

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

PR 893 TTTCTGACTAAGACAGACAAACCGAGCTGTGCTTTGAAGTTA-----GTT 937
QY 56 MetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLys 75
Db 938 GTTATGCGAGCGGTGGATGTTGCCAGGGGTATGGCTTAGTCCATGAG-----CGC 988
QY 76 AlaLeuIleHisArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrVal 95
Db 989 AACTTTATACACCGGATCTAAAGTCAGATAACCTCCTCATATCAGCTGATCGTCCATC 1048
QY 96 LeuLysIleCysAspPheGlyThrAla---CysAspIleGlnThrHis---MetThrAsn 113
Db 1049 ---AAGATTGCTGATTTTGGTGTTCAGAAATGAAGTTCAACCCAGGGATGACACCA 1105
QY 114 AsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu 133
Db 1106 GAGACTGGAACTTACAGATGGATGGCACACAGAGATGCCACAGACACCCCTACACTCAA 1165
QY 134 LysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgLysPro 153
Db 1166 AAAGTGGACGTATAGTTTGGAAATCGTGTGCGAGTTGATTACAGGTCTGTACCG 1225
QY 154 PheAspGluIleGlyGly-----ProAlaPheArgIleMetTrpAlaValHisnGly 171
Db 1226 TTCAGAACATGACCGCGGTTTCAGGCTGCATTTCAGTG-----GTGAACAGAGGA 1276
QY 172 ThrArgProProLeuIleLysAsnLeuProLysProLysProLysLeuMetThrArgCys 191
Db 1277 GTCCGTCCAAACAGTCCACAGCAGATTGTCTTCTGTGCTTGGAGAGATCATGACAGTTGC 1336
QY 192 TrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db 1337 TGGGATGCGGACCCCTGAAAGTCGTCCTTGTGTTTGGAGAGATTGTCATCTTCTG 1390

```

Search completed: December 4, 2003, 09:53:35  
 Job time : 407.459 secs

```

Alignment Scores:
Pred. No.: 1,72e-32 Length: 1631
Score: 373.00 Matches: 87
Percent Similarity: 57.80% Conservative: 39
Best Local Similarity: 39.91% Mismatches: 70
Query Match: 29.79% Indels: 22
DB: 21 Gaps: 9

US-09-830-144-2_COPY_76_303 (1-228) x AAC39537 (1-1631)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 773 GAAGTTTCTATGCTTGTCATTTTGAAGCATCTACATCGTTAGGTTTATTGGTCGTCG 832
QY 22 LeuAsnProVal-----CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 833 ATTAACCGATGGTGTGGTGTGCTGACTGAATATGCAAAAGAGGGGTCTGTCCAGACAG 892
QY 40 ValLeu-----HisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAla 55

```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 07:04:49 ; Search time 17.1149 Seconds  
(without alignments)  
3783.871 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541

Perfect score: 361

Sequence: 1 caaagccgacctaaccct.....agagcgtggtgacagcaccg 204

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09830144/runat\_03122003\_122343\_21194/app\_query.fasta\_1.1230  
-DB=A Geneseq 19Jun03 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 @CGN 1 1.0 @runat\_03122003\_122343\_21194 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DISPATCH=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 19Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	359	99.4	504	18	AAW26706	Human TAB1 (TAK1 b
2	359	99.4	504	18	AAW26707	Human TAB1 (TAK1 b
3	359	99.4	504	20	AAU09541	Human TAB1 protein
4	359	99.4	504	21	AAU91001	Human TAB-1 protei
5	359	99.4	504	21	AAU59450	Human TAB1 protein
6	359	99.4	513	20	AAU09550	Human TAB1 protein
7	359	99.4	516	24	ABU11578	Human MDDT polyep
8	359	99.4	517	20	AAU09546	Human TAB1-FLAG pr
9	352	97.5	84	21	AAU56692	Human prostate can
10	91.5	24.3	125	21	AAU26755	Zea mays protein f
11	88	23.4	1938	24	ABP76682	Streptomyces virid
12	83	22.1	353	24	ABU19816	Androgen-independe
13	82	22.7	1888	22	ABG23214	Novel human diago
14	79.5	21.1	1938	24	ABU98398	Streptomyces virid
15	79.5	22.0	1938	24	ABP76681	Streptomyces virid
16	77.5	20.6	399	24	ABP70654	Recombinant rat fi
17	77	21.3	16	20	AAU09549	Human TAB1 peptide
18	77	20.5	310	22	AAU51391	Propionibacterium
19	76.5	21.2	109	22	AAU39932	Propionibacterium
20	76.5	20.3	146	21	ABG26855	Zea mays protein f
21	76.5	20.3	224	22	ABG16138	Novel human diago
22	76	20.2	266	22	ABG18528	Novel human diago
23	75.5	20.1	689	19	AAU72025	HSV-2 strain SB5 C
24	75	20.8	2038	20	AAU14595	Human T-type volta
25	75	20.8	2044	20	AAU14594	Human T-type volta
26	74.5	19.8	686	22	ABU11397	Human secreted pro
27	74	19.7	127	21	AAU5630	Arabidopsis thalia
28	74	20.5	134	22	AAU63313	Propionibacterium
29	74	19.7	168	21	AAU05629	Arabidopsis thalia
30	74	20.5	205	22	AAU32502	Novel human secret
31	74	19.7	252	21	AAU05628	Arabidopsis thalia
32	74	19.7	1093	14	AAU41001	Human myotonic dys
33	73.5	20.4	130	22	AAU64265	Propionibacterium
34	73.5	20.4	276	22	AAU51816	Propionibacterium
35	73.5	19.5	575	22	ABG21970	Novel human diago
36	73.5	19.5	1938	24	ABP76678	Streptomyces virid
37	73.5	19.5	1938	24	ABP76679	Streptomyces virid
38	73	19.4	1938	24	ABP76681	Streptomyces virid
39	72.5	19.3	201	19	AAU20852	Human neurofilamen
40	72.5	19.3	262	23	ABP69409	Human polypeptide
41	72.5	19.3	386	18	AAU18664	Fragmented human N
42	72.5	20.1	1938	24	ABP76678	Streptomyces virid
43	72	19.1	95	22	AAU45336	Propionibacterium
44	72	19.9	138	22	AAU93357	Human protein sequ
45	72	19.1	495	24	ABP75497	Human secretory po

ALIGNMENTS

RESULT 1  
AAW26706  
ID AAW26706 standard; Protein; 504 AA.  
XX  
AC AAW26706;  
XX  
DT 25-MAR-2003 (updated)  
DT 14-APR-1998 (first entry)  
XX  
Human TAB1 (TAK1 binding protein).  
XX  
TAB1; TAK1 binding protein; transforming growth factor-beta;  
signal transduction; human.  
XX  
Homo sapiens.  
XX  
Key Location/Qualifiers  
FH Misc-difference 52  
FT

/note= "variant has Arg as residue 52"

FT XX EP803571-A2.  
 PN XX 29-OCT-1997.  
 XX PD 24-APR-1997; 97EP-0302808.  
 XX PF 20-NOV-1996; 96US-0752891.  
 XX PR 24-APR-1996; 96JP-0126282.  
 XX PR 28-OCT-1996; 96JP-0300856.  
 XX XX (CHUS ) CHUGAI SEIYAKU KK.  
 XX PI Matsumoto K, Nishida E;  
 XX PI WPI; 1997-515318/48.  
 DR DR N-PSDB; AAT91175.  
 XX XX DNA encoding TAK1 binding protein TAB1 - member of transforming  
 PT growth factor beta receptor signal production pathway, which  
 PT activates TAK-1 kinase activity upon binding  
 XX  
 PS Example 5; Page 17-19; 30pp; English.  
 XX This protein comprises human TAB1, a novel member of the  
 CC transforming growth factor-beta receptor signal transduction  
 CC pathway, which activates TAK-1 kinase activity upon binding. Its  
 CC amino acid sequence was deduced from a cDNA clone (see AAT91175)  
 CC obtained from a kidney library; a variant TAB1 (see AAW26707) has  
 CC Arg rather than Ser at amino acid position 52. Also claimed are:  
 CC isolated DNA encoding a protein modified by a substitution,  
 CC deletion and/or addition of 1 or more amino acids of the 504-residue  
 CC TAB1 sequence; (2) DNA which can hybridise with the 1560 bp TAB1  
 CC nucleic acid sequence; (3) isolated DNA encoding a protein  
 CC comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence;  
 CC (4) DNA encoding a fusion protein comprising an above protein or  
 CC polypeptide; (5) expression vector comprising an above DNA; and  
 CC (6) host cell, preferably a mammalian or yeast cell, transformed by  
 CC the expression vector. Cells expressing TAB1 and TAK1 can be used  
 CC to screen for TGF-beta signalling pathway inhibitors by contacting  
 CC the cells with a test compound, and measuring the TAK1 kinase  
 CC activity.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 504 AA;

Alignment Scores:  
 Pred. No.: 9.39e-34 Length: 504  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.45% Indels: 0  
 DB: 18 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x AAW26706 (1-504)

QY 1 CAAGCCGACCTTAACCTGACGTCCACACACGACGAGAGAGAGTCCAGC 60  
 |||||  
 Db 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 456  
 |||||  
 QY 61 TCTGACGAGGCGCTTCCGCTCCGCCGCCACCTGCTCCGCCGCTGGAGAGCGGT 120  
 |||||  
 Db 457 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 476  
 |||||  
 QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACGCGCTCTGGAGCGTGGACATGGC 180  
 |||||  
 Db 477 ArgValGluProTyrValAspPheAlaGluPheIyrArgLeuTrpSerValAspHisGly 496  
 |||||  
 QY 181 GAGCAGAGCGTGTGTACAGCACCG 204  
 |||||  
 Db 497 GluGlnSerValValThrAlaPro 504

RESULT 2  
 AAW26707  
 ID AAW26707 standard; Protein; 504 AA.  
 XX  
 XX AAW26707;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 14-APR-1998 (first entry)  
 XX  
 XX Human TAB1 (TAK1 binding protein).  
 XX TAB1; TAK1 binding protein; transforming growth factor-beta;  
 KW signal transduction; human.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 52  
 FT /note= "variant has Ser as residue 52"  
 XX  
 XX EP803571-A2.  
 XX 29-OCT-1997.  
 XX PD 24-APR-1997; 97EP-0302808.  
 XX PF 20-NOV-1996; 96US-0752891.  
 XX PR 24-APR-1996; 96JP-0126282.  
 XX PR 28-OCT-1996; 96JP-0300856.  
 XX (CHUS ) CHUGAI SEIYAKU KK.  
 XX Matsumoto K, Nishida E;  
 XX WPI; 1997-515318/48.  
 DR N-PSDB; AAT91178.  
 XX  
 XX DNA encoding TAK1 binding protein TAB1 - member of transforming  
 PT growth factor beta receptor signal production pathway, which  
 PT activates TAK-1 kinase activity upon binding  
 XX  
 PS Example 5; Page 19-21; 30pp; English.  
 XX This protein comprises human TAB1, a novel member of the  
 CC transforming growth factor-beta receptor signal transduction  
 CC pathway, which activates TAK-1 kinase activity upon binding. Its  
 CC amino acid sequence was deduced from a cDNA clone (see AAT91176)  
 CC obtained from a kidney library; a variant TAB1 (see AAW26706) has  
 CC Ser rather than Arg at amino acid position 52. Also claimed are:  
 CC isolated DNA encoding a protein modified by a substitution,  
 CC deletion and/or addition of 1 or more amino acids of the 504-residue  
 CC TAB1 sequence; (2) DNA which can hybridise with the 1560 bp TAB1  
 CC nucleic acid sequence; (3) isolated DNA encoding a protein  
 CC comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence;  
 CC (4) DNA encoding a fusion protein comprising an above protein or  
 CC polypeptide; (5) expression vector comprising an above DNA; and  
 CC (6) host cell, preferably a mammalian or yeast cell, transformed by  
 CC the expression vector. Cells expressing TAB1 and TAK1 can be used  
 CC to screen for TGF-beta signalling pathway inhibitors by contacting  
 CC the cells with a test compound, and measuring the TAK1 kinase  
 CC activity.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 504 AA;

Alignment Scores:  
 Pred. No.: 9.39e-34 Length: 504  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.45% Indels: 0  
 DB: 18 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x AAW26707 (1-504)

QY 1 CAAGAGCCGAGCTTAACCTCAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 60  
 Db 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSer 456  
 QY 61 TCTCAGCAGGAGCCTCTCCGCTCCGCGCCGCCACTCGCTCCGCGCTCCGCGAGGACGGT 120  
 Db 457 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 476  
 QY 121 CGTGTGAGCCCTAARGGACTTTCTGAGTTTACCGCCCTCTGAGCGCTGGACCATGCC 180  
 Db 477 ArgValGluProThrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 496  
 QY 181 GAGCAGAGCGTGGTGACGACCG 204  
 Db 497 GluGlnSerValValThrAlaPro 504

## RESULT 3

AA09541  
 ID AAY09541 standard; Protein; 504 AA.

XX AC AAY09541;  
 XX DT 21-JUL-1999 (first entry)  
 XX DE Human TAB1 protein.  
 XX KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
 KW transforming growth factor beta.  
 XX OS Homo sapiens.  
 XX PN WO9921010-A1.

XX PD 29-APR-1999.

XX PF 22-OCT-1998; 98WO-JP04796.

XX PR 22-OCT-1997; 97JP-0290188.

XX PA (CHUS) CHUGAI SEIYAKU KK.

XX PI Ohtomo T, Ono K, Tsuchiya M;

XX DR WPI; 1999-312645/26.

XX DR N-PSDB; AAX56278.

XX PT Screening for TGF- beta inhibitory substances, which are useful as  
 XX drugs for treatment of diseases relating to its disorder

XX FS Claim 3; Page 147-149; 195pp; Japanese.

XX CC A method has been developed for screening for substances which inhibit  
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method  
 CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 CC and (b) detecting the amount of bound polypeptide, in which the sample  
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming  
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
 CC indications e.g. as TGF-beta signal transmission inhibitors or  
 CC activators, or extracellular matrix protein production enhancement  
 CC inhibitors or activators, or cell proliferation prevention inhibitors or  
 CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence represents human TAB1.

XX SQ Sequence 504 AA;

Alignment Scores:

Pred. No.: 9.39e-34 Length: 504

Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.45% Indels: 0  
 DB: 20 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x AAY09541 (1-504)

QY 1 CAAGAGCCGAGCTTAACCTCAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 60  
 Db 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSer 456  
 QY 61 TCTCAGCAGGAGCCTCTCCGCTCCGCGCCGCCACTCGCTCCGCGCTCCGCGAGGACGGT 120  
 Db 457 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 476  
 QY 121 CGTGTGAGCCCTAARGGACTTTCTGAGTTTACCGCCCTCTGAGCGCTGGACCATGCC 180  
 Db 477 ArgValGluProThrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 496  
 QY 181 GAGCAGAGCGTGGTGACGACCG 204  
 Db 497 GluGlnSerValValThrAlaPro 504

## RESULT 4

AA091001  
 ID AAY91001 standard; Protein; 504 AA.

XX AC AAY91001;  
 XX DT 04-SEP-2000 (first entry)

XX DE Human TAB-1 protein sequence SEQ ID NO:4.  
 XX KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;  
 KW screening; signal transduction; inhibition; inflammatory cytokine;  
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;  
 KW antiinflammatory; suppression.

XX OS Homo sapiens.

XX PN WO200023610-A1.

XX PD 27-APR-2000.

XX PF 21-OCT-1999; 99WO-JP05817.

XX PR 21-OCT-1998; 98JP-0299962.

XX PA (CHUS) CHUGAI SEIYAKU KK.

XX PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;

XX DR WPI; 2000-339707/29.

XX DR N-PSDB; AAA39106.

XX PT Method for screening inhibitors of TAK1 signal transduction for  
 XX suppression of inflammatory cytokine production and use as  
 XX antiinflammatory agents

XX PS Disclosure; Page 90-94; 100pp; Japanese.

XX CC The present invention describes a method for screening compounds for  
 CC inhibition of inflammatory cytokine signal transduction by contacting  
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition  
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds  
 CC for inhibition of inflammatory cytokine signal transduction in which the  
 CC inhibition of TAK1 phosphorylation is selected for; and drug  
 CC compositions for the treatment of inflammatory disorders containing as  
 CC active component an inflammatory cytokine signal transduction inhibitor.  
 CC TAK1 is an essential component of the signalling process which results  
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),  
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used





CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method  
 CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 CC and (b) detecting the amount of bound polypeptide, in which the sample  
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming  
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
 CC indications e.g. as TGF-beta signal transmission inhibitors or  
 CC activators, or extracellular matrix protein production enhancement  
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or  
 CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence represents human TAB1.  
 XX  
 SQ Sequence 513 AA;

Alignment Scores:  
 Pred. No.: 9.41e-34 Length: 513  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.45% Indels: 0  
 DB: 20 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x AAY09550 (1-513)

QY 1 CAAGCCGACCTTAACCTGAGTTCACCAACGACGACGAGCAGTCCAGC 60  
 Db 446 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSer 465  
 QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCCACTCGCTCCGCTCCGCGAGGAGCGT 120  
 Db 466 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 485  
 QY 121 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGCACCATGGC 180  
 Db 486 ArgValGluProTyrValAspPheAlaGluPheIzrArgLeuTrpSerValAspHisGly 505  
 QY 181 GAGCAGAGCGTGGTGACAGCACCG 204  
 Db 506 GluGlnSerValValThrAlaPro 513

RESULT 7  
 ABU11578  
 ID ABU11578 standard; Protein; 516 AA.

XX AC ABU11578;  
 XX  
 DT 12-FEB-2003 (first entry)  
 XX  
 DE Human MDDT polypeptide SEQ ID 525.

XX MDDT; human; disease detection and treatment molecule polypeptide;  
 KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;  
 KW haemostatic; nephrotropic; antianaemic; antiproliferative; hepatotropic;  
 KW gene therapy; protein replacement therapy; cell proliferative disorder;  
 KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma;  
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopenia;  
 KW psoriasis; hepatitis.

XX Homo sapiens.

XX WO200279449-A2.

XX 10-OCT-2002.

XX 27-MAR-2002; 2002WO-US09944.

XX 28-MAR-2001; 2001US-279619P.

XX 29-MAR-2001; 2001US-280067P.

XX 29-MAR-2001; 2001US-280068P.

PR 16-MAY-2001; 2001US-291280P.  
 PR 17-MAY-2001; 2001US-291829P.  
 PR 17-MAY-2001; 2001US-291849P.  
 PR 19-JUN-2001; 2001US-299428P.  
 PR 20-JUN-2001; 2001US-299776P.  
 PR 20-JUN-2001; 2001US-300001P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 PI Daffo GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshay SR;  
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;  
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;  
 XX WPI; 2003-059431/05.  
 DR N-PSDB; ABX34568.

XX New purified disease detection and treatment molecule proteins and  
 PT polynucleotides, useful for diagnosing, treating or preventing cancers  
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis  
 PT or hepatitis  
 XX  
 PS Claim 27; SEQ ID NO 525; 339pp + Sequence Listing; English.

XX This invention describes a novel disease detection and treatment molecule  
 CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,  
 CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,  
 CC antianaemic, antiproliferative and hepatotropic activity. The polynucleotides  
 CC and the polypeptides of the invention can be used for gene therapy,  
 CC protein replacement therapy and are useful for treating a variety of  
 CC diseases or conditions. These polypeptides or polynucleotides are  
 CC particularly useful for diagnosing, treating or preventing cell  
 CC proliferative disorders (e.g. cancers including adenocarcinoma,  
 CC leukaemia, lymphoma, melanoma or sarcoma), anaemia, Crohn's  
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's  
 CC syndromes, inflammation, osteoporosis, thrombocytopenia, psoriasis or  
 CC hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded  
 CC by ABU11450-ABU11845, described in the disclosure of the invention.  
 CC NOTE: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 516 AA;

Alignment Scores:  
 Pred. No.: 9.42e-34 Length: 516  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 99.45% Indels: 0  
 DB: 24 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x ABU11578 (1-516)

QY 1 CAAGCCGACCTTAACCTGAGTTCACCAACGACGACGAGCAGTCCAGC 60  
 Db 449 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSer 468  
 QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCCACTCGCTCCGCTCCGCGAGGAGCGT 120  
 Db 469 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 488  
 QY 121 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGCACCATGGC 180  
 Db 489 ArgValGluProTyrValAspPheAlaGluPheIzrArgLeuTrpSerValAspHisGly 508  
 QY 181 GAGCAGAGCGTGGTGACAGCACCG 204  
 Db 509 GluGlnSerValValThrAlaPro 516

RESULT 8  
 AAY09546

```
ID AAY09546 standard; Protein; 517 AA.
XX
AC AAY09546;
XX
XX 21-JUL-1999 (first entry)
DT
DE Human TAB1-FLAG protein.
XX
XX Human: TAB1; TAK1; screening; inhibition; TGF-beta;
KW transforming growth factor beta.
KW
XX Homo sapiens.
OS
OS Synthetic.
XX
XX WO9921010-A1.
PN
XX 29-APR-1999.
PD
XX 22-OCT-1998; 98WO-JP04796.
PF
XX 22-OCT-1997; 97JP-0290188.
PR
XX (CHUS ) CHUGAI SEIYAKU KK.
PA
XX Ohtomo T, Ono K, Tsuchiya M;
XX
XX WPI; 1999-312645/26.
DR
XX N-PSDB; AAX56282.
DR
XX
XX Screening for TGF- beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX
XX Example 1; Page 163-166; 195pp; Japanese.
XX
XX A method has been developed for screening for substances which inhibit
XX the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX comprises: (a) contacting the polypeptide in the presence of a sample;
XX and (b) detecting the amount of bound polypeptide, in which the sample
XX can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX indications e.g. as TGF-beta signal transduction inhibitors or
XX activators, or extracellular matrix protein production enhancement
XX inhibitors or activators, or cell proliferation prevention inhibitors or
XX activators, or monocyte migration inhibitors or activators, or
XX physiological activity induction inhibitors or activators, or
XX immunosuppression inhibitors or activators, or amyloid beta protein
XX precipitation inhibitors or activators, and such substances can also be
XX inhibitors of the TAK1 polypeptide function, particularly kinase
XX activity. The present sequence represents TAB1-FLAG from an example of
XX the present invention.
XX
XX SQ Sequence 517 AA;

Alignment Scores:
Pred. No.: 9.42e-34 Length: 517
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.45% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-3_COPY_1338_1541 (1-204) x AAY09546 (1-517)

QY 1 CAAGCCGACCTTACCTGAGTCACCAACGACGACGACGACGACGACGACGACGACG 60
Db 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSer 456
QY 61 TCTGACGAGGCGCTCTTCCGCTCCCGCCCGCCGCTCCGCTCCGCTCCGCTCCGCTCCGCT 120
Db 457 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 476
QY 121 CGTGTTCAGCCCTATGTGGACTTCTGTGAGTTTACGCGCTCTGTGGAGCGTGGACCTATGGC 180
|||||
```

477 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 496

181 GAGCAGAGCGTGTGACAGCACCG 204  
|||||

497 GluGlnSerValValThrAlaPro 504

RESULT 9  
AAB56692  
ID AAB56692 standard; Protein; 84 AA.  
XX  
AC AAB56692;  
XX  
XX 13-MAR-2001 (first entry)  
DT  
XX  
DE Human prostate cancer antigen protein sequence SEQ ID NO:1270.  
XX  
XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
KW wound; infectious disease.  
XX  
XX Homo sapiens.  
OS  
XX WO200055174-A1.  
PN  
XX 21-SEP-2000.  
PD  
XX  
XX 08-MAR-2000; 2000WO-US05988.  
PF  
XX 12-MAR-1999; 99US-0124270.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
XX  
XX Rosen CA, Ruben SM;  
XX  
XX WPI; 2000-587513/55.  
DR N-PSDB; AAF15895.  
XX  
XX Prostate cancer associated gene sequences, referred to as prostate  
PT cancer antigens, useful for treatment, prevention, and diagnosis of  
PT disorders such as prostate cancer -  
XX  
XX Claim 11; Page 1693; 2338pp; English.  
XX  
XX AAF15566 to AAF16505 encode the human prostate cancer associated  
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.  
CC The prostate cancer antigens can have neuroprotective, cytostatic,  
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,  
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,  
CC and can be used in gene therapy. The prostate cancer antigen  
CC polynucleotides may be used for detection of prostate cancer, chromosome  
CC identification, as chromosome markers, and for numerous other diagnostic  
CC or research purposes. The prostate cancer antigens may be used to treat  
CC disorders such as neural, immune, muscular, reproductive,  
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to  
CC AAB57303 represent sequences used in the exemplification of the present  
CC invention.  
XX  
XX SQ Sequence 84 AA;

Alignment Scores:  
Pred. No.: 4.8e-33 Length: 84  
Score: 352.00 Matches: 67  
Percent Similarity: 98.53% Conservative: 0  
Best Local Similarity: 98.53% Mismatches: 1  
Query Match: 97.51% Indels: 0  
DB: 21 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x AAB56692 (1-84)

QY	1	CARAGCCGAGCTTAACCTGCAGTCCACCAACGACGACGACGACGACGCTCCAGC	60	PR	28-MAY-1999;
Db	17	GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer	36	PR	01-JUN-1999;
QY	61	TCTCAGCGAGGCGCTCTTCCGGTCCCGGCCCGCCCACTCGCTCCCGCCCTGGCGGACGACGCT	120	PR	03-JUN-1999;
Db	37	Ser***GlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly	56	PR	04-JUN-1999;
QY	121	CGTTTGAGCCCTATGCGACTTTTCTGCTGAGTTTACCGCCCTCTGGAGCGTGGACCATGCG	180	PR	07-JUN-1999;
Db	57	ArgValGluProTyrValaspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly	76	PR	08-JUN-1999;
QY	181	GAGCAGAGCGTGGTGACAGCAGC	204	PR	09-JUN-1999;
Db	77	GluGlnSerValValThrAlaPro	84	PR	10-JUN-1999;
RESULT 10					
AAG26755					
ID	AAG26755	standard; Protein; 125 AA.			
XX	AAG26755;				
XX	XX				
XX	XX				
DT	17-OCT-2000	(first entry)			
DE	Zea mays	protein fragment SEQ ID NO: 31330.			
XX					
KW	Protein identification; signal transduction pathway; metabolic pathway;				
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				
KW	termination sequence; corn.				
XX					
OS	Zea mays	subsp. mays.			
XX					
PN	EP1033405-A2.				
PD	06-SEP-2000.				
XX					
XX					
PF	25-FEB-2000; 2000EP-0301439.				
XX					
PR	25-FEB-1999;	99US-0121825.			
PR	05-MAR-1999;	99US-0123180.			
PR	09-MAR-1999;	99US-0123548.			
PR	23-MAR-1999;	99US-0125788.			
PR	25-MAR-1999;	99US-0126284.			
PR	29-MAR-1999;	99US-0126785.			
PR	01-APR-1999;	99US-0127462.			
PR	06-APR-1999;	99US-0128234.			
PR	08-APR-1999;	99US-0128714.			
PR	16-APR-1999;	99US-0129845.			
PR	19-APR-1999;	99US-0130077.			
PR	21-APR-1999;	99US-0130449.			
PR	23-APR-1999;	99US-0130510.			
PR	23-APR-1999;	99US-0130891.			
PR	28-APR-1999;	99US-0131449.			
PR	30-APR-1999;	99US-0132048.			
PR	04-MAY-1999;	99US-0132407.			
PR	05-MAY-1999;	99US-0132484.			
PR	06-MAY-1999;	99US-0132485.			
PR	06-MAY-1999;	99US-0132486.			
PR	07-MAY-1999;	99US-0132487.			
PR	11-MAY-1999;	99US-0132863.			
PR	14-MAY-1999;	99US-0134256.			
PR	14-MAY-1999;	99US-0134218.			
PR	14-MAY-1999;	99US-0134219.			
PR	18-MAY-1999;	99US-0134370.			
PR	19-MAY-1999;	99US-0134768.			
PR	20-MAY-1999;	99US-0134941.			
PR	21-MAY-1999;	99US-0135124.			
PR	24-MAY-1999;	99US-0135353.			
PR	25-MAY-1999;	99US-0135629.			
PR	27-MAY-1999;	99US-0136021.			
PR	27-MAY-1999;	99US-0136392.			
PR	01-JUN-1999;	99US-0136788.			
PR	03-JUN-1999;	99US-0137222.			
PR	04-JUN-1999;	99US-0137502.			
PR	07-JUN-1999;	99US-0137724.			
PR	08-JUN-1999;	99US-0138094.			
PR	10-JUN-1999;	99US-0138540.			
PR	10-JUN-1999;	99US-0138847.			
PR	14-JUN-1999;	99US-0139119.			
PR	16-JUN-1999;	99US-0139452.			
PR	17-JUN-1999;	99US-0139453.			
PR	17-JUN-1999;	99US-0139492.			
PR	18-JUN-1999;	99US-0139454.			
PR	18-JUN-1999;	99US-0139455.			
PR	18-JUN-1999;	99US-0139456.			
PR	18-JUN-1999;	99US-0139457.			
PR	18-JUN-1999;	99US-0139458.			





PS Claim 20; SEQ ID No 53573; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1888 AA;

Alignment Scores:  
 Pred. No.: 0.596 Length: 1888  
 Score: 82.00 Matches: 27  
 Percent Similarity: 40.70% Conservative: 8  
 Best Local Similarity: 31.40% Mismatches: 27  
 Query Match: 22.71% Indels: 24  
 DB: 22 Gaps: 4

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x ABG23214 (1-1888)

QY 8 CGACCTTAACCTCGAGTCCA-----CCACACGACACACGAGCAGCAGCT 55  
 Db 1794 ArgProProProPheThrGlnThrSerLeuHisProValArgProArgGluGlnGlnAla 1813  
 QY 56 CCAGCTCTGAGGAGGCTCTCGCTCCCGCGCCGCCACTCGCTCCCGCTCGCGAGG 115  
 Db 1814 ProProCysThrSerAlaAlaSerAlaProGlnThrGlnThrArgAlaProThrProArg 1833  
 QY 116 ACG-----GTCTGTTGAGCCCTATGTG----- 139  
 Db 1834 SerProThrArgCysCysArgProLysThrArgThrCysProSerProProTrpSerPro 1853  
 QY 140 -----ACTTTGCTGAGTTTACGCTCTGAGCGTGAGC----- 175  
 Db 1854 SerThrArgThrThrAlaThrCysSerProSerGlyArgTrpThrThrArgProCysArg 1873  
 QY 176 ATGCGCAGCAGCGGTGG 193  
 Db 1874 LeuSerSerAlaAlaTip 1879

RESULT 14  
 ID ABB98398 standard; Protein; 19938 AA.

XX ABB98398;

XX 05-MAR-2003 (first entry)

XX Streptomyces viridochromogenes Avi gene cluster polypeptide frame 1.  
 DE Streptomyces viridochromogenes Avi gene cluster polypeptide frame 1.  
 XX Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;  
 KW medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.  
 XX Streptomyces viridochromogenes.  
 OS Streptomyces viridochromogenes.  
 XX WO200268436-A1.  
 FN WO200268436-A1.  
 XX 06-SEP-2002.

PD 06-SEP-2002.

XX 24-AUG-2001; 2001WO-EP09815.

XX 25-FEB-2001; 2001DE-1009166.

XX (COMB-) COMBINATURE BIOPHARM AG.

XX Weitnauer G, Muehlenweg A, Trefzer A, Bechtold A;

XX WPI; 2003-018650/01.

XX N-PSDB; AB237515.

XX New avilamycin derivatives, useful for treatment of infections, and  
 PT nucleic acid encoding avilamycin synthesis enzymes

XX Example 1; Page 68-301; 319pp; German.

XX The invention relates to avilamycin derivatives (I) with antibacterial,  
 CC virucide, protozoacide and fungicide activity. (I) are useful for  
 CC treatment of infections (bacterial, viral, protozoal or fungal), in human  
 CC or veterinary medicine, particularly where caused by Staphylococcus  
 CC aureus. (I) are more hydrophilic than known avilamycins. The present  
 CC sequence is that of an avilamycin synthesis enzyme from the Streptomyces  
 CC viridochromogenes Avilamycin A biosynthetic gene cluster  
 CC (AB237515-AB237516).

XX Sequence 19938 AA;

Alignment Scores:  
 Pred. No.: 1.65 Length: 19938  
 Score: 79.50 Matches: 21  
 Percent Similarity: 47.62% Conservative: 9  
 Best Local Similarity: 33.33% Mismatches: 20  
 Query Match: 21.14% Indels: 13  
 DB: 24 Gaps: 3

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x ABB98398 (1-19938)

QY 190 CGCTCTGCTCGCATGTCTCCACGCTCCAGAGCGGTAAACTCAGAAAGTCCACATAGG 131  
 Db 18631 ArgProArgArgGlyProArgGlnArgAsnGlyArgThr---ProAsnProArgArg 18649  
 QY 130 GCTCAACACGACCTCTCTCGCCAGCG-----GGAGCGAGTGGCGCGCGGAGC 80  
 Db 18650 ProArgArgArgProArgSerAlaProGlyArgArgAlaGlyArgGlnValGlyArg 18669  
 QY 79 GGAAGAGCGC-----CTCCGTCAGAGCTGGAGCTGCTGC 47  
 Db 18670 GlyArgAlaAsnAlaArgGlnSerArgLeuGlyArgValArgAlaGlyProLeuCys 18689

QY 46 TCTCGGTGT 38

Db 18690 AsnSerCys 18692

RESULT 15

ABP76681

ID ABP76681 standard; Protein; 19938 AA.

XX ABP76681;

XX 26-FEB-2003 (first entry)

XX Streptomyces viridochromogenes Avi gene cluster polypeptide frame 5.

XX Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;  
 KW medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.

XX Streptomyces viridochromogenes.

OS WO200268436-A1.

XX 06-SEP-2002.

```
XX 24-AUG-2001; 2001WO-EP09815.
XX
XX 25-FEB-2001; 2001DE-1009166.
XX
XX (COMB-) COMBINATURE BIOPHARM AG.
XX
XX Weitnauer G, Muehlenweg A, Trefzer A, Bechthold A;
XX
XX MPI; 2003-018650/01.
XX
XX N-PSDB; ABZ37516.
XX
XX New avilamycin derivatives, useful for treatment of infections, and
XX nucleic acid encoding avilamycin synthesis enzymes
XX
XX Example 1; Page 68-301; 319pp; German.
XX
XX The invention relates to avilamycin derivatives (I) with antibacterial,
XX virucide, protozoacide and fungicide activity. (I) are useful for
XX treatment of infections (bacterial, viral, protozoal or fungal), in human
XX or veterinary medicine, particularly where caused by Staphylococcus
XX aureus. (I) are more hydrophilic than known avilamycins. The present
XX sequence is that of an avilamycin synthesis enzyme from the Streptomyces
XX viridochromogenes Avilamycin A biosynthetic gene cluster
XX (ABZ37515-ABZ37516).
XX
XX SQ Sequence 19938 AA;

Alignment Scores:
Pred. No.: 1.65 Length: 19938
Score: 79.50 Matches: 29
Percent Similarity: 37.50% Conservative: 1
Best Local Similarity: 36.25% Mismatches: 24
Query Match: 22.02% Indels: 26
DB: 24 Gaps: 5

US-09-830-144-3_COPY_1338_1541 (1-204) x ABP76681 (1-19938)
QY 8 CGACCTTAACCTGC-----AGTCACCAACACGACGAGAGCAGCAGCT 55
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2075 ArgPro---ProCysGlyThrProTrpSerProProArgArgAlaArgArgTrpPro 2093
QY 56 CCAGCTCTGACGGAGGCTCTTCGCTCCGGC-----CGCCCCACTCGCTCCCGC 106
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2094 ProProAlaArgAlaGlyGlySerAlaProGlyCysGlyThrProArgArgGlySer-Pr 2113
QY 107 CTGG-----CGAGGACGGTCGTGTTGAGCCCTATGTGGACTTTG 145
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2113 oTrpAlaTyArgSerGlyThrAlaArgGlyArgSerAla----- 2126
QY 146 CTGAGTTTACCGCTCTGGAGCGTGACCATGGCGAGCAGCGTGTGACGACC 203
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
2127 -----ProProThrArgArgArgProTrpSerGlyGlySerGlyAspSerSer 2142
```

Search completed: December 4, 2003, 09:02:56  
Job time : 29.1149 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: December 4, 2003, 09:00:20 ; Search time 6.2027 Seconds  
(without alignments)  
2783.114 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541  
Perfect score: 361  
Sequence: 1 caaagccgacctaaccct.....agaggtgtgacagcaccg 204

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310859 residues  
Total number of hits satisfying chosen parameters: 657434  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_n2p.model -DEV=xlh  
-Q/cgn2\_1/USPTO spool/US09830144/runat\_03122003\_122345\_21337/app query.fasta\_1.1230  
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPL=0  
-COOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 -CCN\_1\_1\_29 @runat\_03122003\_122345\_21337 -NCPU=6 -ICPU=3  
-NO MAP -LARGEOUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pcp.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pcp.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pcp.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pcp.\*  
5: /cgn2\_6/prodata/1/iaa/6C COMB.pcp.\*  
6: /cgn2\_6/prodata/1/iaa/6D COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359	99.4	504	2	US-08-752-891-2
2	359	99.4	504	2	US-08-752-891-6
3	359	99.4	504	2	US-09-144-178-2
4	359	99.4	504	2	US-09-144-178-6
5	359	99.4	504	3	US-09-406-854-2
6	359	99.4	504	3	US-09-406-854-6
7	359	99.4	504	4	US-09-529-279-2
8	359	99.4	504	4	US-10-158-895-2
9	359	99.4	513	4	US-09-529-279-43
10	359	99.4	513	4	US-10-158-895-43
11	359	99.4	517	4	US-09-529-279-11
12	359	99.4	517	4	US-10-158-895-11

13	100.5	27.8	136	4	US-09-252-991A-22734	Sequence 22734, A
14	92.5	24.6	189	4	US-09-252-991A-17056	Sequence 17056, A
15	89.5	23.8	303	4	US-09-252-991A-19799	Sequence 19799, A
16	86.5	23.0	496	4	US-09-252-991A-22592	Sequence 22592, A
17	84.5	22.5	363	4	US-09-252-991A-20195	Sequence 20195, A
18	84	22.3	333	4	US-09-252-991A-24478	Sequence 24478, A
19	84	22.3	721	4	US-09-252-991A-28293	Sequence 28293, A
20	83.5	22.2	338	4	US-09-252-991A-18820	Sequence 18820, A
21	83.5	22.2	480	4	US-09-252-991A-20125	Sequence 20125, A
22	83	22.1	264	4	US-09-252-991A-24662	Sequence 24662, A
23	82	21.8	269	4	US-09-252-991A-26605	Sequence 26605, A
24	81.5	21.7	200	4	US-09-252-991A-19793	Sequence 19793, A
25	81.5	21.7	355	4	US-09-252-991A-23226	Sequence 23226, A
26	81	21.5	201	4	US-09-252-991A-25745	Sequence 25745, A
27	81	21.5	1228	4	US-09-252-991A-17764	Sequence 17764, A
28	80.5	21.4	239	4	US-09-252-991A-25387	Sequence 25387, A
29	80.5	21.4	589	4	US-09-252-991A-32631	Sequence 32631, A
30	80	21.3	338	4	US-09-252-991A-19609	Sequence 19609, A
31	79.5	22.0	136	4	US-09-252-991A-23572	Sequence 23572, A
32	79	21.0	243	4	US-09-252-991A-26015	Sequence 26015, A
33	79	21.0	350	4	US-09-252-991A-19537	Sequence 19537, A
34	78.5	20.9	268	4	US-09-252-991A-27950	Sequence 27950, A
35	78.5	20.9	297	4	US-09-252-991A-18932	Sequence 18932, A
36	78	20.7	416	4	US-09-252-991A-32875	Sequence 32875, A
37	78	20.7	710	4	US-09-252-991A-24946	Sequence 24946, A
38	77.5	20.6	246	4	US-09-252-991A-23345	Sequence 23345, A
39	77.5	20.6	345	4	US-09-252-991A-32325	Sequence 32325, A
40	77	21.3	16	4	US-09-529-279-41	Sequence 41, Appl
41	77	21.3	16	4	US-10-158-895-41	Sequence 41, Appl
42	77	20.5	279	4	US-09-252-991A-26860	Sequence 26860, A
43	77	20.5	586	4	US-09-252-991A-28601	Sequence 28601, A
44	77	20.5	763	4	US-09-252-991A-30146	Sequence 30146, A
45	76.5	21.2	224	4	US-09-252-991A-19251	Sequence 19251, A

ALIGNMENTS

RESULT 1  
US-08-752-891-2  
; Sequence 2, Application US/0875891  
; Patent No. 5837819  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 504 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-752-891-2

Alignment Scores:  
Pred. No.: 2,61e-34 Length: 504  
Score: 359.00 Matches: 68  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 99.45% Indels: 0  
DB: 2 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x US-08-752-891-2 (1-504)

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACGACGACGACGAGGAGCTCCAGC 60  
Db 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSer 456  
QY 61 TCTGACGAGGAGCTCTCCGCTCCGCGCCGCCACTCGCTCCGCGGAGGAGCGT 120  
Db 457 SeraspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluaspGly 476  
QY 121 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCTCTGAGGAGTGGACCATGGC 180  
Db 477 ArgValGluProTyrValaspPheAlaGluPheTyrArgLeuTrpSerValaspHisGly 496  
QY 181 GAGCAGAGCGTGTGACGACCG 204  
Db 497 GluGlnSerValValThrAlaPro 504

RESULT 2

US-08-752-891-6  
Sequence 6, Application US/08752891  
Patent No. 5837819  
GENERAL INFORMATION:  
APPLICANT: MATSUMOTO, Kunihiro  
APPLICANT: NISHIDA, Eisuke  
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/752,891  
FILING DATE: 20-NOV-1996  
CLASSIFICATION: 415  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:

102e

TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 504 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-752-891-6

Alignment Scores:  
Pred. No.: 2,61e-34 Length: 504  
Score: 359.00 Matches: 68  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 99.45% Indels: 0  
DB: 2 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x US-08-752-891-6 (1-504)

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACGACGACGACGAGGAGCTCCAGC 60  
Db 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSer 456  
QY 61 TCTGACGAGGAGCTCTCCGCTCCGCGCCGCCACTCGCTCCGCGGAGGAGCGT 120  
Db 457 SeraspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluaspGly 476  
QY 121 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCTCTGAGGAGTGGACCATGGC 180  
Db 477 ArgValGluProTyrValaspPheAlaGluPheTyrArgLeuTrpSerValaspHisGly 496  
QY 181 GAGCAGAGCGTGTGACGACCG 204  
Db 497 GluGlnSerValValThrAlaPro 504

RESULT 3

US-09-144-178-2  
Sequence 2, Application US/09144178  
Patent No. 5989862  
GENERAL INFORMATION:  
APPLICANT: MATSUMOTO, Kunihiro  
APPLICANT: NISHIDA, Eisuke  
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/144,178  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/752,891  
FILING DATE: 20-NOV-1996  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768

9

102e





```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-3_COPY_1338_1541 (1-204) x US-09-529-279-2 (1-504)
QY 1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGACGAGCAGCTCCAGC 60
DB 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 456
QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCCACCTCGCTCCGCTCGCGAGGACGGT 120
DB 457 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 476
QY 121 CGTGTGAGCCCTATGAGACTTGTGAGTTTACCGCTCTCGAGCGTGGACCATGGC 180
DB 477 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 496
QY 181 GAGCAGAGCGTGTGACAGCACC 204
DB 497 GluGlnSerValThrAlaPro 504

RESULT 8
US-10-158-895-2
; Sequence 2, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-2

Alignment Scores:
Pred. No.: 2,61e-34 Length: 504
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-3_COPY_1338_1541 (1-204) x US-10-158-895-2 (1-504)
QY 1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGACGAGCAGCTCCAGC 60
DB 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 456
QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCCACCTCGCTCCGCTCGCGAGGACGGT 120
DB 457 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 476
QY 121 CGTGTGAGCCCTATGAGACTTGTGAGTTTACCGCTCTCGAGCGTGGACCATGGC 180
DB 477 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 496
QY 181 GAGCAGAGCGTGTGACAGCACC 204
DB 497 GluGlnSerValThrAlaPro 504

RESULT 9
US-09-529-279-43
; Sequence 43, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-43

Alignment Scores:
Pred. No.: 2,62e-34 Length: 513
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-3_COPY_1338_1541 (1-204) x US-09-529-279-43 (1-513)
QY 1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGACGAGCAGCTCCAGC 60
DB 446 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 465
QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCCACCTCGCTCCGCTCGCGAGGACGGT 120
DB 466 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 485
QY 121 CGTGTGAGCCCTATGAGACTTGTGAGTTTACCGCTCTCGAGCGTGGACCATGGC 180
DB 486 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 505
QY 181 GAGCAGAGCGTGTGACAGCACC 204
DB 506 GluGlnSerValThrAlaPro 513

RESULT 10
US-10-158-895-43
; Sequence 43, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-43

Alignment Scores:
Pred. No.: 2,62e-34 Length: 513
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-3_COPY_1338_1541 (1-204) x US-10-158-895-43 (1-513)
QY 1 CAAGCCGACCTTAACCTGAGTCCACACACGACGAGGAGCTCCAGC 60
Db 446 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 465
QY 61 TCTGACGAGGCTCTTCCTGCTCCCGCCCGCCACTCTCCGCTCCGCGGAGGACGGT 120
Db 466 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 485
QY 121 CGTGTTGAGCCCTATGTCAGTTTCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180
Db 486 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuThrSerValAspHisGly 505
QY 181 GAGCAGAGCGTGTGACACACCG 204
Db 506 GluGlnSerValThrAlaPro 513

RESULT 11
US-09-529-279-11
; Sequence 11, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: TOSHIO, TOSHIHIKO
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-11

Alignment Scores:
Pred. No.: 2,63e-34 Length: 517
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-3_COPY_1338_1541 (1-204) x US-10-158-895-11 (1-517)
QY 1 CAAGCCGACCTTAACCTGAGTCCACACACGACGAGGAGCTCCAGC 60
Db 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 456
QY 61 TCTGACGAGGCTCTTCCTGCTCCCGCCCGCCACTCTCCGCTCCGCGGAGGACGGT 120
Db 457 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 476
QY 121 CGTGTTGAGCCCTATGTCAGTTTCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180
Db 477 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuThrSerValAspHisGly 496

RESULT 12
US-10-158-895-11
; Sequence 11, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TOSHIO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-11

Alignment Scores:
Pred. No.: 2,63e-34 Length: 517
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.45% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-3_COPY_1338_1541 (1-204) x US-10-158-895-11 (1-517)
QY 1 CAAGCCGACCTTAACCTGAGTCCACACACGACGAGGAGCTCCAGC 60
Db 437 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 456
QY 61 TCTGACGAGGCTCTTCCTGCTCCCGCCCGCCACTCTCCGCTCCGCGGAGGACGGT 120
Db 457 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 476
QY 121 CGTGTTGAGCCCTATGTCAGTTTCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180
Db 477 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuThrSerValAspHisGly 496

RESULT 13
US-09-252-991A-22734
; Sequence 22734, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
```

```
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22734
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22734

Alignment Scores:
Pred. No.: 0.00696 Length: 136
Score: 100.50 Matches: 27
Percent Similarity: 55.93% Conservative: 6
Best Local Similarity: 45.76% Mismatches: 23
Query Match: 27.84% Indels: 3
DB: 4 Gaps: 1

US-09-830-144-3_COPY_1338_1541 (1-204) x US-09-252-991A-22734 (1-136)
QY 23 AGTCCACCAACACGACACGACGAGCAGCTCCAGCTCTGACGAGGCTCTTCGCT 82
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
55 AlaProProArgThrArgArgSerAlaAlaSerAlaThrSerAsnSerCysSerThr 74
QY 83 CCGCGCGCGCCACTCGCTCCGCTCGGCGAGGAGCGTGTGTGAGCCCTATGTGACT 142
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
75 SerAlaProProThrArgSerArgThrProThrThrArgThrSerSerProSerSerThr 94
QY 143 TTGCTGAGTTTACCGCTCTGGA-----GCGTGACCATGGCGACGAGCG 190
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
95 GlyCysSerSerThrProSerSerSerAlaProAlaThrSerThrSerSerArgAla 113

RESULT 14
US-09-252-991A-17056
; Sequence 17056, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17056
; LENGTH: 189
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17056

Alignment Scores:
Pred. No.: 0.00667 Length: 189
Score: 92.50 Matches: 26
Percent Similarity: 41.33% Conservative: 5
Best Local Similarity: 34.67% Mismatches: 25
Query Match: 24.60% Indels: 19
DB: 4 Gaps: 2

US-09-830-144-3_COPY_1338_1541 (1-204) x US-09-252-991A-17056 (1-189)
QY 193 CCAGGCTCTGCTCCGCTGTCACGCTCCAGAGCGGTAAACTCAGCAAACTCCACAT 134
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
20 ProArgArgProArgHisArgProGlyGlnTrpProGlyArgArgProProHis 39
QY 133 AGGCTCAACACGACCGCT----- 116
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
40 ArgGlyTrpGlnAspArgHisArgArgAlaGlyArgLeuGlnArgProGluAsp 59
QY 115 CCTCGCGCGCGGAGCGAGTGGCGGCGGCGGAGCGGAGAGAGCGCTCCGCTCAGAGCTGG 56
Db :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
60 ProArgArgArgProGlyGlyArgAlaGlyThrGlyArgSerGluArgArg----- 77

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22734
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19799
; Sequence 19799, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19799
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19799

Alignment Scores:
Pred. No.: 0.017 Length: 303
Score: 89.50 Matches: 28
Percent Similarity: 35.58% Conservative: 9
Best Local Similarity: 26.92% Mismatches: 28
Query Match: 23.80% Indels: 39
DB: 4 Gaps: 5

US-09-830-144-3_COPY_1338_1541 (1-204) x US-09-252-991A-19799 (1-303)
QY 199 CTCTCACCAGCTCTGCTCCGCTCCGCTCCAGCTCCAGAGCGGTAAA----- 152
Db :||| :||| :||| :||| :||| :||| :||| :|||
181 ValGlyProGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 200
QY 151 -----ACTCAGCAAACTCCACATAGGCTCAACACGACCGCTCCTCCGCGGCG-- 104
Db :||| :||| :||| :||| :||| :||| :||| :|||
201 ProAlaAspLeuArgArgAlaProHisArgArgHisGlyProAlaArgGlyValPro 220
QY 103 GGAGCGAGTGGCGGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 68
Db :||| :||| :||| :||| :||| :||| :||| :|||
221 GlyGlyAlaGlyArgThrAlaLeuGlyArgProGlyAspGlyGlyGlyAlaAspGlyHis 240
QY 67 CGTCAGAGC-----TGAGGCTGC 50
Db :||| :||| :||| :||| :||| :||| :||| :|||
241 ArgArgSerProCysArgLeuLeuProSerSerGlyAlaCysProAlaThrTrpArgCys 260
QY 49 -----TGCTCTGCGTGTGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 17
Db :||| :||| :||| :||| :||| :||| :||| :|||
261 GlySerSerSerProAsnSerAlaCysSerAlaCysSerAlaCysCysTrpArgSerProAla 280
QY 16 TTAAGGTCGGGC 5
Db :||| :||| :||| :||| :||| :||| :||| :|||
281 ProArgProGly 284

Search completed: December 4, 2003, 09:09:43
Job time : 8.2027 secs
```

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 09:07:25 ; Search time 36.2973 Seconds  
(without alignments)  
2090.550 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541  
Perfect score: 361  
Sequence: 1 caaagccgacctaaccot.....agagcgtggacagcaccg 204

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 1368560

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DSV=xlh  
-Q=/cgn2\_1/USFTO.spool/US09830144/runat\_03122003\_122347\_21456/app\_query.fasta\_1.1230  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09830144 -CGCN\_1\_1\_78 @runat\_03122003\_122347\_21456  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=120 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description

DR

1	359	99.4	504	12	US-10-384-743-2	Sequence 2, Appli
2	359	99.4	504	14	US-10-123-427-2	Sequence 2, Appli
3	359	99.4	504	14	US-10-123-427-6	Sequence 6, Appli
4	359	99.4	504	14	US-10-158-895-2	Sequence 2, Appli
5	359	99.4	513	12	US-10-384-743-43	Sequence 43, Appli
6	359	99.4	513	14	US-10-158-895-43	Sequence 43, Appli
7	359	99.4	517	12	US-10-384-743-11	Sequence 11, Appli
8	359	99.4	517	14	US-10-158-895-11	Sequence 11, Appli
9	352	27.5	84	10	US-09-925-300-1270	Sequence 1270, Ap
c 10	83	22.1	3534	15	US-10-184-644-417	Sequence 417, App
c 11	83	22.1	3534	15	US-10-184-634-417	Sequence 417, App
c 12	77	21.3	16	12	US-10-384-743-41	Sequence 41, Appli
c 13	77	21.3	16	14	US-10-158-895-41	Sequence 41, Appli
c 14	74.5	19.8	613	12	US-10-260-937-16	Sequence 16, Appli
c 15	74	19.9	252	16	US-10-278-536-180	Sequence 180, App
c 16	72	19.9	574	12	US-10-168-097A-76	Sequence 76, Appli
c 17	72	19.9	574	12	US-10-239-431A-38	Sequence 38, Appli
c 18	72	19.1	644	12	US-10-063-685-55	Sequence 55, Appli
c 19	72	19.1	644	15	US-10-184-644-213	Sequence 213, App
c 20	72	19.1	644	15	US-10-184-634-213	Sequence 213, App
c 21	71	18.9	323	9	US-09-925-301-952	Sequence 952, App
c 22	71	19.7	505	12	US-10-442-017-17	Sequence 17, Appli
c 23	71	18.9	2916	12	US-10-140-472-69	Sequence 69, Appli
c 24	71	18.9	2916	12	US-10-141-761-69	Sequence 69, Appli
c 25	71	18.9	2916	12	US-10-142-885-69	Sequence 69, Appli
c 26	71	18.9	2916	12	US-10-158-790-69	Sequence 69, Appli
c 27	71	18.9	2916	12	US-10-137-871-69	Sequence 69, Appli
c 28	71	18.9	2916	12	US-10-140-805-69	Sequence 69, Appli
c 29	71	18.9	2916	12	US-10-140-864-69	Sequence 69, Appli
c 30	71	18.9	2916	12	US-10-140-923-69	Sequence 69, Appli
c 31	71	18.9	2916	12	US-10-141-756-69	Sequence 69, Appli
c 32	71	18.9	2916	12	US-10-141-759-69	Sequence 69, Appli
c 33	71	18.9	2916	15	US-10-123-153-69	Sequence 69, Appli
c 34	71	18.9	2916	16	US-10-146-731-69	Sequence 69, Appli
c 35	70.5	19.5	55	9	US-09-864-761-38194	Sequence 38194, A
c 36	70.5	18.8	4679	9	US-09-804-898-2	Sequence 2, Appli
c 37	70	19.4	573	15	US-10-156-761-11836	Sequence 11836, A
c 38	69.5	18.5	105	9	US-09-867-550-1874	Sequence 1874, Ap
c 39	69.5	19.3	199	10	US-09-941-831-21	Sequence 21, Appli
c 40	69.5	18.5	478	15	US-10-156-761-11945	Sequence 11945, A
c 41	69	18.4	967	12	US-10-017-161-1766	Sequence 1766, Ap
c 42	69	18.4	3442	12	US-10-140-472-235	Sequence 235, App
c 43	69	18.4	3442	12	US-10-141-761-235	Sequence 235, App
c 44	69	18.4	3442	12	US-10-142-885-235	Sequence 235, App
c 45	69	18.4	3442	12	US-10-158-790-235	Sequence 235, App

ALIGNMENTS

RESULT 1  
US-10-384-743-2  
; Sequence 2, Application US/10384743  
; Publication No. US20030162228A1  
; GENERAL INFORMATION:

; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/Jp98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: Jp 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 504  
; TYPE: PRT  
; ORGANISM: Homo sapiens







QY 181 GAGCAGCGCTGTGACAGCACC 204  
Db |||||||  
497 GluGlnSerValValThrAlaPro 504

## RESULT 9

US-09-925-300-1270  
; Sequence 1270, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; APPLICANT: Steve Ruben,  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1270  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (38)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-300-1270

## Alignment Scores:

Pred. No.: 6,97e-30 Length: 84  
Score: 352.00 Matches: 67  
Percent Similarity: 98.53% Conservative: 0  
Best Local Similarity: 98.53% Mismatches: 1  
Query Match: 97.51% Indels: 0  
DB: 10 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x US-09-925-300-1270 (1-84)

QY 1 CAAGCCGACCTTACCTGCGAGTCCACCAACGACAGCAGCAGCTCCAGC 60  
Db |||||||  
17 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSer 36  
QY 61 TCTGACGAGGCGCTCTCCGCTCCCGCCCGCCACTCGCTCCGCTGCGGAGGAGGT 120  
Db |||||||  
37 Ser\*\*\*GlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 56  
QY 121 CGTGTGAGCCCTATGCTGAGCTTTGCTAGTTTACCGCTCTGGAGCGTGGACCATGGC 180  
Db |||||||  
57 ArgValGluProTyValaspPheAlaGluPheTyArgLeuTrpSerValaspHisGly 76  
QY 181 GAGCAGCGCTGTGACAGCACC 204  
Db |||||||  
77 GluGlnSerValValThrAlaPro 84

## RESULT 10

US-10-184-644-417  
; Sequence 417, Application US/10184644  
; Publication No. US2003004930A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3430R1C227  
; CURRENT APPLICATION NUMBER: US/10/184,644  
; PRIOR FILING DATE: 2002-06-28  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 417  
; LENGTH: 3534  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-184-644-417

## Alignment Scores:

Pred. No.: 1.95 Length: 3534  
Score: 83.00 Matches: 16  
Percent Similarity: 61.76% Conservative: 5  
Best Local Similarity: 47.06% Mismatches: 13  
Query Match: 22.07% Indels: 0  
DB: 15 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x US-10-184-644-417 (1-3534)

QY 106 GCGGAGCGAGTGGCGGCGGAGCGAGAGGCGCTCCGTCAGAGCTGGAGCTGCTGC 47  
Db |||||||  
2654 AlaAlaAlaThrGlyCysAlaGlyThrGlyThrGlyAlaCysThrThrThrCysCys 2673  
QY 46 TCTGCGTGTGCGTGTGGTGGAGTGCAGCGGTTAAGGTTCGGGC 5  
Db |||||||  
2674 CysAlaCysAlaCysAlaCysThrGlyGlyAlaThrThrGly 2687

## RESULT 11

US-10-184-634-417  
; Sequence 417, Application US/10184634  
; Publication No. US2003006868A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430R1C217  
; CURRENT APPLICATION NUMBER: US/10/184,634  
; CURRENT FILING DATE: 2002-06-28  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 417  
; LENGTH: 3534  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-184-634-417

## Alignment Scores:

Pred. No.: 1.95 Length: 3534  
Score: 83.00 Matches: 16  
Percent Similarity: 61.76% Conservative: 5  
Best Local Similarity: 47.06% Mismatches: 13  
Query Match: 22.07% Indels: 0  
DB: 15 Gaps: 0

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x US-10-184-634-417 (1-3534)

QY 106 GCGGAGCGAGTGGCGGCGGAGCGAGAGGCGCTCCGTCAGAGCTGGAGCTGCTGC 47  
Db |||||||  
2654 AlaAlaAlaThrGlyCysAlaGlyThrGlyThrGlyAlaCysThrThrThrCysCys 2673





***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 07:06:54 ; Search time 4.82432 Seconds  
(without alignments)  
3977.116 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541

Perfect score: 361

Sequence: 1 caaagccgacctaacct.....agagcgtgtgacagcaccg 204

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09830144/runat\_03122003\_122344\_21226/app\_query.fasta\_1.1230  
-DB=SwissProt 41 -QFWT=fastan -SUFFIX=rsrp -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 @CGN 1 1 16 @runat\_03122003\_122344\_21226 -NCFU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359	99.4	504	TAB1_HUMAN	Q15750 homo sapien
2	73	20.2	518	VL2_HPV47	P22425 human papill
3	72	19.9	1199	P121_RAT	P52591 rattus norv
4	71	19.7	132	VE4_HPV2A	P25483 human papill
5	71	19.7	519	VL2_HPV14	P36749 human papill
6	71	19.7	521	VL2_HPV49	P36762 human papill
7	71	19.7	578	DSBD_PANCI	Q9x2b2 pantoea cic
8	70.5	19.5	518	VL2_HPV12	P36748 human papill
9	70	19.4	198	GSCI_MOUSE	P56916 mus musculu
10	70	19.4	518	VL2_HPV20	P50794 human papill
11	69	19.4	520	VL2_HPV19	P36752 human papill
12	69	19.1	895	GNDS_RAT	Q03386 rattus norv
13	68.5	18.2	239	RGSK_MOUSE	Q9qzbl mus musculu
14	68.5	18.2	537	VL2B_RAT	P26896 rattus norv
15	68.5	18.2	824	TGM1_RAT	P26896 rattus norv
16	68	18.1	361	IE63_PVKA	Q85232 pseudorabie
17	66	17.6	99	NICI_HUMAN	Q9ug19 homo sapien
18	66	18.3	427	TEA4_MOUSE	Q62296 mus musculu

## ALIGNMENTS

### RESULT 1

TAB1\_HUMAN STANDARD; PRT; 504 AA.

AC Q15750;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Mitogen-activated protein kinase kinase 7 interacting protein 1

DE (TAK1-binding protein 1).

GN MAP3K7IP1 OR TAB1.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OC NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=96216294; PubMed=8639164;

RA Shibuya H., Yamaguchi K., Shirakabe K., Tonegawa A., Gotoh Y.,

RA Ueno N., Irie K., Nishida E., Matsumoto K.;

RT "TAB1: an activator of the TAK1 MAPKKK in TGF-beta signal

transduction.";

RL Science 272:1179-1182 (1996).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=20057165; PubMed=10591208;

RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,

RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,

RA Baggeley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,

RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,

RA Burdill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,

RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,

RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,

RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,

RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,

RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,

RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,

RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,

RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,

RA Mcclellan J., McLaren S., McMurray A., Milne S.A., Mortimore B.J.,

RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,

Q9uk39 homo sapien  
P06918 human papill  
P06419 human papill  
P50827 human papill  
P22735 homo sapien  
Q94966 homo sapien  
P50798 human papill  
O00716 homo sapien  
P26540 human papill  
Q9p123 homo sapien  
P24107 human immun  
Q9sw09 arabidopsis  
Q22053 caenorhabdi  
P29376 homo sapien  
P47806 mus musculu  
Q9cx84 mus musculu  
O70521 rattus norv  
P31263 notophthalm  
P22909 rattus norv  
Q9bxq6 homo sapien  
P54817 bovine herp  
Q91su1 oryza sativ  
Q9sp35 arabidopsis  
O19032 orycteropus  
P78325 homo sapien  
P34684 caenorhabdi  
Q01338 mus musculu

102b (5)

QY	121	CGTGTGAGCCCTATGTGGA	CTTTGTGTGAGTTTACCGCTCTGGAGCGTGGACCATGGC	180
Db	477	ArgValGluProTyrValAsp	PheAlaGluPheTyrArgLeuTrpSerValAspHisGly	496
QY	181	GAGCAGAGCGTGTGTGAC	GAGCACCGC	204
Db	497	GluGlnSerValValThr	AlaPro	504
RESULT 2				
VL2	HPV47	STANDARD;	PRT;	518 AA.
ID	VL2 HPV47			
AC	P22425;			
DT	01-AUG-1991 (Rel. 19, Created)			
DT	01-AUG-1991 (Rel. 19, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	Minor capsid protein L2.			
GN	L2.			
OS	Human papillomavirus type 47.			
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			
OC	Papillomavirus.			
OX	NCBI_TaxID=10594;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=90281611; PubMed=2162112;			
RA	Kiyono T., Adachi A., Ishibashi M.;			
RT	"Genome organization and taxonomic position of human papillomavirus			
RT	type 47 inferred from its DNA sequence.";			
RL	Virology 177:401-405(1990).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announcement">http://www.isb-sib.ch/announcement</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M32305; AAA46981.1; -			
DR	PIR; F35324; P2WL47.			
DR	InterPro; IPR000784; Late_L2.			
DR	Pfam; PF00513; Late_protein_L2; 1.			
KW	Coat protein; Late protein.			
SQ	SEQUENCE 518 AA; 56436 MW; 1E99381FF640BF0 CRC64;			
Alignment Scores:				
Pred. No.:	4.17	Length:	518	
Score:	73.00	Matches:	22	
Percent Similarity:	49.21%	Conservative:	9	
Best Local Similarity:	34.92%	Mismatches:	26	
Query Match:	20.22%	Indels:	6	
DB:	1	Gaps:	2	
US-09-830-144-3_COPY_1338_1541 (1-204) x VL2_HP47 (1-518)				
QY	11	CCCTTAACCCCTGAGTCCA	-----CCACACGACGACGAGCAGCAGCTCCA	58
Db	141	ProLleProGluGlyProThrIleAspSerProValValThrThrThrGlySerSer	160	
QY	59	GCTCTGACGAGGCGCTCTTCGCTCCGCGCCCGCCACTCGCTCCGCTCGCGAGGACG	118	
Db	161	AlaValLeuGluValAlaProGluProValProProThrArgValArgIleAlaArgThr	180	
QY	119	GTCGTGTGAGCCCTATGTGGACTTGTCTG	-----AGTTTACCGCCTCTCGGAGCGTG	172
Db	181	GlnTyrHisAsnProSerPheGlnIleLeuThrGluSerThrProAlaGlnGlyGluSer	200	
QY	173	ACCATGGCG	181	
Db	201	SerLeuAla	203	
RESULT 3				
ID	P121_RAT	STANDARD;	PRT;	1199 AA.
ID	P121_RAT			

```

QY 137 TGGACTTTGCTGAGTTTACCGCTCTGAGCGTGGACCATGCGGAGAGA 187
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 GlyThrAlaThrSerThrValAlaThrGlyThrThrAlaSerAlaSerLys 801

RESULT 4
ID_VE4_HPV2A STANDARD; PRT; 132 AA.
AC P25483;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE Probable E4 protein.
OS Human papillomavirus type 2a.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10584;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91188699; PubMed=1964523;
RA Hirsch-Behnam A.; Delius H.; de Villiers E.M.;
RT "A comparative sequence analysis of two human papillomavirus (HPV)
RT types 2a and 57."
RL Virus Res. 18:81-98 (1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X55964; -; NOT_ANNOTATED_CDS.
DR PIR; S15618; S15618.
DR InterPro; IPR003861; Papilloma_E4.
DR Pfam; PF02711; Pap_E4; 1.
KW Early protein.
SQ SEQUENCE 132 AA; 14532 MW; 8256491E0DES19C1 CRC64;

Alignment Scores:
Pred. No.: 6.2 Length: 132
Score: 71.00 Matches: 22
Percent Similarity: 45.45% Conservative: 3
Best Local Similarity: 40.00% Mismatches: 20
Query Match: 19.67% Indels: 10
DB: Gaps: 3

US-09-830-144-3_COPY1_338_1541 (1-204) x VE4_HPV2A (1-132)

QY 11 CTTTAACCTCGAGTCCACCA-----ACAGCACAGCAGACGAGCTCGAGCTCTG 64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 ProlysA:G:CysAlaProArgGlnArgValArgArgProSerAlaSerValSer 85
      :::::
QY 65 ACGGAGGCTCTTCGCTCCGCGCCGCCCACTGCTCCCGCTGCGGACGACGTCGTG 124
      :::::
Db 86 SerSerAspSerSerIleProGlyProThrLeuArgGluSerGluArg----- 102
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 125 TTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCTCTTGAGCG 169
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 103 -----GlyLysTrp-----SerValThrSerGlyAla 112

RESULT 5
VL2_HPV14
ID_VL2_HPV14 STANDARD; PRT; 519 AA.
AC P36749;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Minor capsid protein L2.
GN GN L2.
OS Human papillomavirus type 14.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.

```

```

OX NCBI_TaxID=10605;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74467; CAA52504.1; -.
DR PIR; S36471; S36471;
DR InterPro; IPR000784; Late L2.
DR Pfam; PF00513; late protein L2; 1.
KW Coat protein; Late protein.
SQ SEQUENCE 519 AA; 56873 MW; 52E21620474CD1E8 CRC64;

Alignment Scores:
Pred. No.: 6.83 Length: 519
Score: 71.00 Matches: 16
Percent Similarity: 50.00% Conservative: 11
Best Local Similarity: 29.63% Mismatches: 23
Query Match: 19.67% Indels: 4
DB: 1 Gaps: 1

US-09-830-144-3_COPY_1338_1541 (1-204) x VL2_HPV14 (1-519)

QY 2 AAGCCCGCACTTAACCTGCAGTCCACCAACA-----GGCACAGCAGACA 49
Db 138 GluValHisProGlyProSerArgProProThrAspThrProValThrSerThrGly 157
QY 50 GCAGCTCCAGCTCAGCAGGAGGCTTCGCTCCGCGCGCGCGCGCTGCTCCGCGCTG 109
Db 158 GlySerSerAlaIleLeuGluValAlaProGluProProProSerArgValArgVal 177
QY 110 GCAGCAGCTGCTGTGTGACCCATGTCGACTTGTGCTGAGT 151
Db 178 ThrArgThrGlnTyrHisAsnProSerPheGlnValIleThr 191

RESULT 6
VL2_HPV49
ID_VL2_HPV49 STANDARD; PRT; 521 AA.
AC P36762;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Minor capsid protein L2.
GN L2.
OS Human papillomavirus type 49.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10616;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74467; CAA52504.1; -.
DR PIR; S36471; S36471;
DR InterPro; IPR000784; Late L2.
DR Pfam; PF00513; late protein L2; 1.
KW Coat protein; Late protein.
SQ SEQUENCE 519 AA; 56873 MW; 52E21620474CD1E8 CRC64;

Alignment Scores:
Pred. No.: 6.83 Length: 521
Score: 71.00 Matches: 16
Percent Similarity: 50.00% Conservative: 11
Best Local Similarity: 29.63% Mismatches: 23
Query Match: 19.67% Indels: 4
DB: 1 Gaps: 1

US-09-830-144-3_COPY_1338_1541 (1-204) x VL2_HPV49 (1-521)

QY 38 ACACGACGACGACGACGCTCCAGCTCAGCGAGGCTTCGCTCCGCGCGCGCGCT 97
Db 153 ThrThrSerArgGlySerSerAlaValLeuGluValAlaSerGluProThrProThr 172
QY 98 CGCTCCGCTCGCAGGACGCTGCTGTGAGCCCTATGTCGACTTGTGCTGAGT 151
Db 173 ArgThrArgIleSerArgThrGlnTyrHisAsnProSerPheGlnIleLeuThr 190

RESULT 7
DSBD_PANCI
ID_DSBD_PANCI STANDARD; PRT; 578 AA.
AC Q9XDB2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thiol:disulfide interchange protein dsbD precursor (EC 1.8.1.8)
DE (Protein-disulfide reductase) (Disulfide reductase).
GN DSBD.
OS Pantoea citrea.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pantoea.
OX NCBI_TaxID=53336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1056R;
RX MEDLINE=20200361; PubMed=10735866;
RA Fujol C.J., Kado C.I.;
RT "Genetic and biochemical characterization of the pathway in Pantoea
citrea leading to pink disease of pineapple.";
RL J. Bacteriol. 182:2230-2237(2000).
CC -!- FUNCTION: Required to facilitate the formation of correct
CC disulfide bonds in some periplasmic proteins and for the assembly
CC of the periplasmic c-type cytochromes. Acts by transferring
CC electrons from cytoplasmic thioredoxin to the periplasm. This
CC transfer involves a cascade of disulfide bond formation and
CC reduction steps (By similarity).
CC -!- CATALYTIC ACTIVITY: Protein dithiol + NAD(P)+ = protein disulfide
CC + NAD(P)H.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (By similarity).
CC -!- SIMILARITY: BELONGS TO THE THIOREDOXIN FAMILY. DSBD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF102175; AAD38449.1; -.
DR HAMAP; MF_00399; 1
DR InterPro; IPR003834; Cytococh_TM.
DR InterPro; IPR006662; Thioredo.
DR Pfam; PF02683; DsbD; 1.

```

DR PROSITE; PS00194; THIOREDOXIN; 1.  
 KW Oxidoreductase; Redox-active center; Electron transport; NAD;  
 FT Transmembrane; Inner membrane; Cytochrome c-type biogenesis; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 578  
 FT DOMAIN 23 180  
 FT TRANSMEM 181 201  
 FT DOMAIN 202 216  
 FT TRANSMEM 217 237  
 FT DOMAIN 238 253  
 FT TRANSMEM 254 274  
 FT DOMAIN 275 306  
 FT TRANSMEM 307 327  
 FT DOMAIN 328 336  
 FT TRANSMEM 337 357  
 FT DOMAIN 358 367  
 FT TRANSMEM 368 388  
 FT DOMAIN 389 394  
 FT TRANSMEM 395 415  
 FT DOMAIN 416 428  
 FT TRANSMEM 429 449  
 FT DOMAIN 450 578  
 FT DISULFID 125 131  
 FT DISULFID 193 315  
 FT DISULFID 492 495  
 SQ SEQUENCE 578 AA; 62547 MW; 383B32797B7FA0B4 CRC64;

Alignment Scores:  
 Pred. No.: 6.89 Length: 578  
 Score: 71.00 Matches: 23  
 Percent Similarity: 45.61% Conservations: 3  
 Best Local Similarity: 40.35% Mismatches: 13  
 Query Match: 19.67% Indels: 18  
 DB: 1 Gaps: 4

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x DSBD\_PANCI (1-578)  
 QY 20 TGCAGTCCACCA---ACAGC-----ACAGCAGAGCAGCA 52  
 Db 131 CysTyProProGluThrArgGluValProLeuSerGlnValSerThrArgSerGlu 150  
 QY 53 GTCGAGCTGTACGAGGAGCTCTCCGCTCCGCGCCG-----CCCACT 97  
 Db 151 AlaProAlaThrAlaAlaThrProAlaProValProGluProGlnSerGlyProAla 170  
 QY 98 CGCTCCGCGCTCGAGGAGGCTGCTGTTGAGCCCTATGTGGACTTGTGTG 148  
 Db 171 ValSerArgLeuPro-----PheSerProLeuTyrAlaLeuLeu 183

RESULT 8  
 VL2 HPV12  
 ID VL2 HPV12 STANDARD; PRT; 518 AA.  
 AC P36748;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Minor capsid protein L2.  
 GN L2.  
 OS Human papillomavirus type 12.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10604;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94265501; PubMed=8205838;  
 RA Dellus H., Hofmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types.";  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; X74466; CAAS2500.1; -.  
 DR PIR; S36542; S36542.  
 DR InterPro; IPR000784; Late\_L2.  
 DR Pfam; PF00513; late protein L2; 1.  
 KW Coat protein; late protein.  
 SQ SEQUENCE 518 AA; 56776 MW; EFDFODB9BED22A9B CRC64;

## Alignment Scores:

Pred. No.: 7.73 Length: 518  
 Score: 70.50 Matches: 20  
 Percent Similarity: 56.36% Conservations: 11  
 Best Local Similarity: 36.36% Mismatches: 21  
 Query Match: 19.53% Indels: 3  
 DB: 1 Gaps: 2

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x VL2 HPV12 (1-518)  
 QY 23 AGTCCACCAACACGACGACGACGAGCTCCAGCTCTGAGGAGGCTCTTCGCT 82  
 Db 150 SerProValValThrThrSerArgGlySerSer---AlaLeuGluValAlaProAsp 169  
 QY 83 CCGCGCCGCGCCACTCGCTCCGCTCCGCTGCGGAGGAGCGTGTGTGAGCCCTATGTGGACT 142  
 Db 169 ProIleProProThrArgValAlaArgThrGlnTyrHisAsnProAlaPheGln 188  
 QY 143 TTGCTG-----AGTTTACCGCTCTGGAGCGTGACCATGGCG 181  
 Db 189 IleIleThrGluSerThrProAlaGlnGlyGluThrSerLeuAla 203

## RESULT 9

GSCL MOUSE  
 ID GSCL MOUSE STANDARD; PRT; 198 AA.  
 AC P56916;

DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Homeobox protein goosecoid-like (GSC-2).  
 GN GSCL  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97189340; PubMed=9037598;  
 RA Gallili N., Baldwin H.S., Lund J., Reeves R., Gong W., Wang Z.,  
 RA Roe B.A., Emanuel B.S., Nayak S., Mickanin C., Budarf M.L., Buck C.A.;  
 RT "A region of mouse chromosome 16 is syntenic to the DiGeorge,  
 RT velocardiiofacial syndrome minimal critical region.";  
 RL Genome Res. 7:17-26(1997).  
 RN [2]  
 RP ERRATUM.  
 RX MEDLINE=97264346; PubMed=9110179;  
 RA Gallili N., Baldwin H.S., Lund J., Reeves R., Gong W., Wang Z.,  
 RA Roe B.A., Emanuel B.S., Nayak S., Mickanin C., Budarf M.L., Buck C.A.;  
 RL Genome Res. 7:399-399(1997).  
 RN [3]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=98110571; PubMed=9441739;  
 RA Funke B., St Jore B., Puech A., Sirotkin H., Edelman L., Carlson C.,  
 RA Raft S., Pandita R.K., Kucherlapati R., Skoultschi A., Morrow B.E.;  
 RT "Characterization and mutation analysis of goosecoid-like (GSCL), a  
 RT homeodomain-containing gene that maps to the critical region for  
 RT VCFS/DGS on 22q11.";  
 RL Genomics 46:364-372(1997).

CC -!- FUNCTION: MAY HAVE A ROLE IN DEVELOPMENT. MAY REGULATE ITS OWN  
 CC -!- TRANSCRIPTION. MAY BIND THE BICOID CONSENSUS SEQUENCE TAATCC.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT TESTIS.



DR	SMART; SMO0147; RasGEF; 1.
DR	SMART; SMO0229; RasGFEN; 1.
DR	PROSITE; PS00720; RASGEF; 1.
DR	PROSITE; PS50200; RA; 1.
DR	PROSITE; PS50009; RASGEF_CAT; 1.
DR	PROSITE; PS50212; RASGEF_NTER; 1.
KW	Guanine-nucleotide releasing factor; 3D-structure.
FT	DOMAIN 112 237
FT	FT DOMAIN 367 629
FT	FT DOMAIN 779 866
FT	FT STRAND 779 787
FT	FT STRAND 796 801
FT	TURN 802 803
FT	STRAND 805 805
FT	FT HELIX 806 815
FT	TURN 816 818
FT	HELIX 824 826
FT	STRAND 827 834
FT	TURN 835 836
FT	STRAND 837 840
FT	TURN 843 844
FT	STRAND 846 846
FT	HELIX 847 850
FT	TURN 853 854
FT	STRAND 858 863
SQ	SEQUENCE 895 AA; 98869 MW; 43E1674675A4E1C9 CRC64;

Alignment Scores:

Pred. No.:	11.7	Length:	895
Score:	69.00	Matches:	14
Percent Similarity:	51.43%	Conservative:	4
Best Local Similarity:	40.00%	Mismatches:	9
Query Match:	19.11%	Indels:	8
DB:	1	Gaps:	1

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x GNDS\_RAT (1-895)

QY	6	CCCGACCTTAACCTTCAGTCCACCACACGACGACGACGACGACTCCAGCTCAT
Ddb	321	ProGlueLeuApproThrValSerGlnSerLeuHisIeuGlu-----
QY	66	CGGAGGCCTCTCCGCTCCCGGCCGGCCCACTCGCTCCCGCTGG 110
Ddb	335	-----ProAlaProValProAlaProAlaIeuGluProSerTrp 347

RESULT 13

RGSK\_MOUSE

IAD	RGSK_MOUSE	STANDARD;	PRT;	239 AA.
OC	Q9ZB1; Q9CUV8; Q9OZB2;			
DDT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DDT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Regulator of G-protein signaling 20 (RGS20) (Regulator of G-protein signaling Z1).			
DE	RGS20 OR RGS21.			
OOS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	[1]			
RRN	SEQUENCE FROM N.A., AND VARIANT ARG-32.			
RC	STRAIN=129/B6, and BALB/c;			
RC	Barker S.A., Wang J., Ross E.M.;			
RT	"A mouse ortholog of RGS21."			
RLN	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.			
RL	[2]			
PC	SEQUENCE OF 5-239 FROM N.A.			
XP	STRAIN=C57BL/6J; TISSUE=Hippocampus;			
XX	MEDLINE=21085660; PubMed=11217851;			
AA	Kawai J., Shingagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
AA	Arakawa T., Haya A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
AA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,			
AA	Saito T., Okazaki Y., Gotchieri T., Bere H., Kashimura			





```

DB: 1 Gaps: 1
US-09-830-144-3_COPY_1338_1541 (1-204) x IL2B_RAT (1-537)
QY 131 GGCTCAACACGACCGCTCTCCGACGCGGAGCGAGTGGCGGCGGAGCGGAGAGG 72
Db 216 GlyLysThrArgThrTrpSerProTrpSerGlnProMetAlaPheArgThrArgProAla 235
QY 71 CCTCCGTCAGAGCTG----- 57
Db 236 AspProLysGluIlePheProLeuProTrpLeuArgCysLeuLeuValLeuGlyCys 255
QY 56 -----GAGCTGCTGCTCTGCGTGTGCTGTGTGTGGTGGACTGCAGG 18
Db 256 PhePheGlyPheLeuSerCysValCysValLeuValLysCysArg 270

RESULT 15
TGM1_RAT
ID TGM1_RAT STANDARD; PRT; 824 AA.
AC P23606;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Protein-glutamine gamma-glutamyltransferase K (EC 2.3.2.13)
DE (Transglutaminase K) (TGase K) (TGK) (TG(K)) (Transglutaminase 1)
DE (Epidermal TGase).
GN TGM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=91067700; PubMed=1979171;
RA Phillips M.A., Stewart B.E., Qin Q., Chakravarty R., Floyd E.E.,
RA Jetten A.M., Rice R.H.;
RT "Primary structure of keratinocyte transglutaminase.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:9333-9337(1990).
CC -1- FUNCTION: Catalyzes the cross-linking of proteins and the
CC conjugation of polyamines to proteins. Responsible for cross-
CC linking epidermal proteins during formation of the stratum
CC corneum.
CC -1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N(5)-
CC alkylglutamine + NH(3).
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Membrane-bound.
CC -1- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M57263; AAA63495.1; --
CC PIR; B38423; B38423.
CC HSP; P00488; IQRK.
CC InterPro; IPR001102; GluttransfG.
CC InterPro; IPR002931; Trnsglutase like.
CC Pfam; PF01841; Transglut_core; 1.
CC Pfam; PF00927; Transglutamin_C; 2.
CC Pfam; PF00868; Transglutamin_N; 1.
CC SMART; SM00460; TGC; 1.
CC PROSITE; PS00547; TRANSGLUTAMINASES; 1.
CC Transference; Acyltransferase; Calcium-binding; Membrane;
KW Keratinization.
FT ACT_SITE 385 385 BY SIMILARITY.
FT ACT_SITE 444 444 BY SIMILARITY.
FT ACT_SITE 467 467 BY SIMILARITY.
FT METAL 507 507 CALCIUM (BY SIMILARITY).
FT METAL 509 509 CALCIUM (BY SIMILARITY).

```

```

FT METAL 556 556 CALCIUM (BY SIMILARITY).
FT METAL 561 561 CALCIUM (BY SIMILARITY).
SQ SEQUENCE 824 AA; 90769 MW; A7D81C148CEFD938 CRC64;

Alignment Scores:
Pred. No.: 13.1 Length: 824
Score: 68.50 Matches: 18
Percent Similarity: 40.35% Conservative: 5
Best Local Similarity: 31.58% Mismatches: 21
Query Match: 18.22% Indels: 13
DB: 1 Gaps: 3

US-09-830-144-3_COPY_1338_1541 (1-204) x TGM1_RAT (1-824)
QY 175 GGTCCAGCTCCAGAGCGCGTAAA-----ACTCAGCAAGT 140
Db 3 GlyProArgSerAspValGlyArgTrpGlyArgSerProTrpGlnProThrProSer 22
QY 139 CCACATAGGCGCTCAACACGACCGCTCTCGCCAGCGGAGCGAGTGGCGGCGGAGC 80
Db 23 ProGluProGluProGluProGluProGluProGluProGluProGluProGluProGlu 42
QY 79 GGAAGAGCGCTCCGTCACAGCTGGAGC---TGCTGCTCTGCGTGTGCGTGT 32
Db 43 -----ArgSerPheTrpAlaArgCysCysGlyCysCysSerCys 55

Search completed: December 4, 2003, 09:03:42
Job time : 7.82432 secs

```

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 00:51:52 ; Search time 944.878 Seconds  
(without alignments) 8332.421 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541

Perfect score: 204

Sequence: 1 caaagccgcgaccttaaccct.....agagcgtggtagacagcaccg 204

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	1515	6	E31041 Method for
2	204	100.0	1560	6	AR058299 Sequence
3	204	100.0	1560	6	AR058302 Sequence
4	204	100.0	1560	6	AR088273 Sequence
5	204	100.0	1560	6	AR088276 Sequence
6	204	100.0	1560	6	AR116881 Sequence
7	204	100.0	1560	6	AR116884 Sequence
8	204	100.0	1560	6	AR231191 Sequence
9	204	100.0	1560	6	AR307972 Sequence
10	204	100.0	1560	6	AR307972 Sequence
11	204	100.0	1568	6	AR307972 Sequence
12	204	100.0	1568	6	AR308004 Sequence
13	204	100.0	1569	6	AR231195 Sequence
14	204	100.0	1569	6	AR307976 Sequence
15	204	100.0	3096	9	HSU49928
16	204	100.0	3253	9	BC050554
17	204	100.0	3290	9	BC038582
18	204	100.0	6960	9	HS407F17
19	204	100.0	17118	2	AC141444
20	164.6	80.7	1784	10	AB088136
21	164.6	80.7	2943	10	BC041110
22	164.6	80.7	2944	10	BC027054
23	164.6	80.7	182516	2	AC140267
24	159.8	78.3	288805	2	AC127784
25	159.8	78.3	298235	2	AC127784
26	96.8	47.5	1926	5	XLU92031
27	40	19.6	276800	1	SC0939115
28	37.4	18.3	53121	2	AC136331
29	37.4	18.3	181545	10	AL626774
30	37.2	18.2	1485	6	E15521
31	37.2	18.2	1491	6	AR073849
32	37.2	18.2	1491	9	BT006992
33	37.2	18.2	1491	12	BT007934
34	37.2	18.2	1505	6	E16553
35	37.2	18.2	1648	6	E04413
36	37.2	18.2	1653	9	HUMAAT
37	37.2	18.2	1871	9	BC018207
38	37.2	18.2	10622	1	AE004819
39	37.2	18.2	36792	3	LMFL3747
40	36.2	17.7	299800	1	AP005028
41	36	17.6	301700	1	AP005948
42	35.8	17.5	41630	9	AC005785
43	35.8	17.5	57075	2	AC110746
44	35.8	17.5	162090	10	AL805921
45	35.8	17.5	174678	9	AL450332

# ALIGNMENTS

RESULT 1  
E31041  
LOCUS E31041 Method for screening substance inhibiting binding to XIAP. PAT 18-JUN-2001  
DEFINITION E31041  
ACCESSION E31041  
VERSION E31041.1 GI:13017306  
KEYWORDS JP 199326328-A/1.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1515)  
AUTHORS Kunihiro,M.  
TITLE Method for screening substance inhibiting binding to XIAP  
JOURNAL Patent: JP 199326328-A 1 26-NOV-1999;  
KUNIHIRO MATSUMOTO

[illegible]

ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 1560)
TITLE	Matsuomoto,K. and Nishida,E.
JOURNAL	TAB1 protein and DNA coding therefore
FEATURES	Patent: US 6140042-A 1 31-OCT-2000;
source	Location/Qualifiers
	1..1560
BASE COUNT	332 a 469 c 480 g 279 t
ORIGIN	/organism="unknown"
Query Match	100.0%; Score 204; DB 6; Length 1560;
Best Local Similarity	100.0%; Pred. No. 1.6e-37;
Matches 204; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CAAGGCCGACCTTAAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 60
Db	1338 CAAGGCCGACCTTAAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1397
QY	61 TCTGACGGAGGCGCTCTTCGCTCCCGGCCCGCCACTCGTCCCGCTCGCGAGGACGGT 120
Db	1398 TCTGACGGAGGCGCTCTTCGCTCCCGGCCCGCCACTCGTCCCGCTCGCGAGGACGGT 1457
QY	121 CGTGTGTAGCGCTATGTGACCTTGTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180
Db	1458 CGTGTGTAGCGCTATGTGACCTTGTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1517
QY	181 GAGCAGAGCGTGGTGACAGCACCG 204
Db	1518 GAGCAGAGCGTGGTGACAGCACCG 1541
RESULT 7	
AR116884	
LOCUS	AR116884 1560 bp DNA linear PAT 16-MAY-2001
DEFINITION	Sequence 5 from patent US 6140042.
ACCESSION	AR116884
VERSION	AR116884.1 GI:14097790
KEYWORDS	Unknown.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 1560)
AUTHORS	Matsuomoto,K. and Nishida,E.
TITLE	TAB1 protein and DNA coding therefore
JOURNAL	Patent: US 6140042-A 5 31-OCT-2000;
FEATURES	Location/Qualifiers
source	1..1560
BASE COUNT	333 a 468 c 480 g 279 t
ORIGIN	/organism="unknown"
Query Match	100.0%; Score 204; DB 6; Length 1560;
Best Local Similarity	100.0%; Pred. No. 1.6e-37;
Matches 204; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CAAGGCCGACCTTAAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 60
Db	1338 CAAGGCCGACCTTAAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1397
QY	61 TCTGACGGAGGCGCTCTTCGCTCCCGGCCCGCCACTCGTCCCGCTCGCGAGGACGGT 120
Db	1398 TCTGACGGAGGCGCTCTTCGCTCCCGGCCCGCCACTCGTCCCGCTCGCGAGGACGGT 1457
QY	121 CGTGTGTAGCGCTATGTGACCTTGTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180
Db	1458 CGTGTGTAGCGCTATGTGACCTTGTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1517
QY	181 GAGCAGAGCGTGGTGACAGCACCG 204
Db	1518 GAGCAGAGCGTGGTGACAGCACCG 1541

	PAT 20-DEC-2002
RESULT 8	
AR231191	DNA linear
LOCUS	1560 bp
DEFINITION	Sequence 1 from patent US 6451617.
ACCESSION	AR231191
VERSION	AR231191.1 GI:27272079
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	Ono,K., Ohtomo,T. and Tsuchiya,M.
TITLE	Method of screening TGF-beta inhibitory substances
JOURNAL	Patent: US 6451617-A 1 17-SEP-2002;
FEATURES	Location/Qualifiers 1..1560 /source= "unknown"
BASE COUNT	332 a 469 c 480 g 279 t
ORIGIN	
Query Match	100.0%; Score 204; DB 6; Length 1560;
Best Local Similarity	100.0%; Pred.No.1.6e-37;
Matches 204; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CAAGGCCGCGACTTAACTTCGGCTCCGCGCCGCACCACACGACGACGACGACGAGCAGCTCCAGC 60
Db	1338 CAAAAGCCCAGCTTAACCTTGCAATCCCAAACGACGACGACGACGAGCATGTCCAGC 1397
Qy	61 TC TGACGGAGGCGCTTTCCGGCTCCGCGCCGCACCACCTGCTCCCGCTCGCGAGGACGGT 120
Db	1398 CTCGACGGAGGCGCTTTCCGGCTCCGCGCCGCACCACCTGCTCCCGCTCGCGAGGACGGT 1457
Qy	121 CGTGTTGAGCGCCTATTGTGGAATTTTGCCTGAGTTTTAACCGCTCTGAGCGGTGGACCATGGC 180
Db	1458 CGTGTTGAGCGCCTATTGTGGAATTTTGCCTGAGTTTTAACCGCTCTGAGCGGTGGACCATGGC 1517
Qy	181 GAGCAGAGCGGTGGTGACGACCG 204
Db	1518 GAGCAGAGCGGTGGTGACGACCG 1541
RESULT 9	PAT 12-JUN-2003
AR307972	DNA linear
LOCUS	1560 bp
DEFINITION	Sequence 1 from patent US 6551840.
ACCESSION	AR307972
VERSION	AR307972.1 GI:31698729
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unclassified.
REFERENCE	1 (bases 1 to 1560)
AUTHORS	Ono,K., Ohtomo,T. and Tsuchiya,M.
TITLE	Method of screening TGF-beta-inhibiting substances
JOURNAL	Patent: US 6551840-A 1 22-APR-2003;
FEATURES	Location/Qualifiers 1..1560 /source= "unknown"
BASE COUNT	332 a 469 c 480 g 279 t
ORIGIN	
Query Match	100.0%; Score 204; DB 6; Length 1560;
Best Local Similarity	100.0%; Pred.No.1.6e-37;
Matches 204; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CAAGGCCGCGACTTAACTTCGGCTCCGCGCCGCACCACGACGACGACGACGAGCAGCTCCAGC 60
Db	1338 CAAAAGCCCAGCTTAACCTTGCAATCCCAAACGACGACGACGACGAGCATGTCCAGC 1397
Qy	61 TC TGACGGAGGCGCTTTCCGGCTCCGCGCCGCACCACCTGCTCCCGCTCGCGAGGACGGT 120
Db	1398 CTCGACGGAGGCGCTTTCCGGCTCCGCGCCGCACCACCTGCTCCCGCTCGCGAGGACGGT 1457

Db 1518 GAGCAGAGCGTGTGACAGCAGCG 1541  
|||||  
RESULT 11  
AR2311223  
LOCUS 1568 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 42 from patent US 6451617.  
ACCESSION AR2311223  
VERSION AR2311223.1 GI:27272111  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Ono,K., Ohmoto,T. and Tsuchiya,M.  
TITLE Method of screening TGF-beta. inhibitory substances  
JOURNAL Patent: US 6451617-A 42 17-SEP-2002;  
FEATURES  
source 1..1568  
/organism="unknown"  
BASE COUNT 339 a 472 c 477 g 280 t  
ORIGIN  
Query Match 100.0%; Score 204; DB 6; Length 1568;  
Best Local Similarity 100.0%; Pred. No. 1.6e-37;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGCCGACCTTAACCTCTGACGTCCACCAACGACGACGACGACGACGCTCCAGC 60  
Db 1346 CAAAGCCGACCTTAACCTCTGACGTCCACCAACGACGACGACGACGACGCTCCAGC 1405  
QY 61 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCGCTCCGCGTGGCGAGGACGGT 120  
Db 1406 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCGCTCCGCGTGGCGAGGACGGT 1465  
QY 121 CGTGTGAGCCCTATGTCGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180  
Db 1466 CGTGTGAGCCCTATGTCGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1525  
QY 181 GAGCAGAGCGTGTGACAGCAGCG 204  
Db 1526 GAGCAGAGCGTGTGACAGCAGCG 1549  
RESULT 12  
AR308004  
LOCUS 1568 bp DNA linear PAT 12-JUN-2003  
DEFINITION Sequence 42 from patent US 6551840.  
ACCESSION AR308004  
VERSION AR308004.1 GI:31698761  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Ono,K., Ohmoto,T. and Tsuchiya,M.  
TITLE Method of screening TGF-beta. inhibiting substances  
JOURNAL Patent: US 6551840-A 42 22-APR-2003;  
FEATURES  
source 1..1568  
/organism="unknown"  
BASE COUNT 339 a 472 c 477 g 280 t  
ORIGIN  
Query Match 100.0%; Score 204; DB 6; Length 1568;  
Best Local Similarity 100.0%; Pred. No. 1.6e-37;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGCCGACCTTAACCTCTGACGTCCACCAACGACGACGACGACGACGCTCCAGC 60  
Db 1346 CAAAGCCGACCTTAACCTCTGACGTCCACCAACGACGACGACGACGACGCTCCAGC 1405  
RESULT 13  
AR2311195  
LOCUS 1569 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 10 from patent US 6451617.  
ACCESSION AR2311195  
VERSION AR2311195.1 GI:27272083  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1569)  
AUTHORS Ono,K., Ohmoto,T. and Tsuchiya,M.  
TITLE Method of screening TGF-beta. inhibitory substances  
JOURNAL Patent: US 6451617-A 10 17-SEP-2002;  
FEATURES  
source 1..1569  
/organism="unknown"  
BASE COUNT 343 a 466 c 476 g 284 t  
ORIGIN  
Query Match 100.0%; Score 204; DB 6; Length 1569;  
Best Local Similarity 100.0%; Pred. No. 1.6e-37;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGCCGACCTTAACCTCTGACGTCCACCAACGACGACGACGACGACGCTCCAGC 60  
Db 1315 CAAAGCCGACCTTAACCTCTGACGTCCACCAACGACGACGACGACGACGCTCCAGC 1374  
QY 61 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCGCTCCGCGTGGCGAGGACGGT 120  
Db 1375 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCGCTCCGCGTGGCGAGGACGGT 1434  
QY 121 CGTGTGAGCCCTATGTCGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180  
Db 1435 CGTGTGAGCCCTATGTCGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1494  
QY 181 GAGCAGAGCGTGTGACAGCAGCG 204  
Db 1495 GAGCAGAGCGTGTGACAGCAGCG 1518  
RESULT 14  
AR307976  
LOCUS 1569 bp DNA linear PAT 12-JUN-2003  
DEFINITION Sequence 10 from patent US 6551840.  
ACCESSION AR307976  
VERSION AR307976.1 GI:31698733  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1569)  
AUTHORS Ono,K., Ohmoto,T. and Tsuchiya,M.  
TITLE Method of screening TGF-beta. inhibiting substances  
JOURNAL Patent: US 6551840-A 10 22-APR-2003;  
FEATURES  
source 1..1569  
/organism="unknown"  
BASE COUNT 343 a 466 c 476 g 284 t  
ORIGIN

QY 61 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCGCTCCGCGTGGCGAGGACGGT 120  
Db 1406 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCGCTCCGCGTGGCGAGGACGGT 1465  
QY 121 CGTGTGAGCCCTATGTCGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180  
Db 1466 CGTGTGAGCCCTATGTCGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1525  
QY 181 GAGCAGAGCGTGTGACAGCAGCG 204  
Db 1526 GAGCAGAGCGTGTGACAGCAGCG 1549  
RESULT 13  
AR2311195  
LOCUS 1569 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 10 from patent US 6451617.  
ACCESSION AR2311195  
VERSION AR2311195.1 GI:27272083  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1569)  
AUTHORS Ono,K., Ohmoto,T. and Tsuchiya,M.  
TITLE Method of screening TGF-beta. inhibitory substances  
JOURNAL Patent: US 6451617-A 10 17-SEP-2002;  
FEATURES  
source 1..1569  
/organism="unknown"  
BASE COUNT 343 a 466 c 476 g 284 t  
ORIGIN  
Query Match 100.0%; Score 204; DB 6; Length 1569;  
Best Local Similarity 100.0%; Pred. No. 1.6e-37;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGCCGACCTTAACCTCTGACGTCCACCAACGACGACGACGACGACGCTCCAGC 60  
Db 1315 CAAAGCCGACCTTAACCTCTGACGTCCACCAACGACGACGACGACGACGCTCCAGC 1374  
QY 61 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCGCTCCGCGTGGCGAGGACGGT 120  
Db 1375 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCGCCACTCGCTCCGCGTGGCGAGGACGGT 1434  
QY 121 CGTGTGAGCCCTATGTCGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180  
Db 1435 CGTGTGAGCCCTATGTCGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1494  
QY 181 GAGCAGAGCGTGTGACAGCAGCG 204  
Db 1495 GAGCAGAGCGTGTGACAGCAGCG 1518  
RESULT 14  
AR307976  
LOCUS 1569 bp DNA linear PAT 12-JUN-2003  
DEFINITION Sequence 10 from patent US 6551840.  
ACCESSION AR307976  
VERSION AR307976.1 GI:31698733  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1569)  
AUTHORS Ono,K., Ohmoto,T. and Tsuchiya,M.  
TITLE Method of screening TGF-beta. inhibiting substances  
JOURNAL Patent: US 6551840-A 10 22-APR-2003;  
FEATURES  
source 1..1569  
/organism="unknown"  
BASE COUNT 343 a 466 c 476 g 284 t  
ORIGIN

Query Match 100.0%; Score 204; DB 6; Length 1569;  
Best Local Similarity 100.0%; Pred. No. 1.6e-37;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 60  
DB 1315 CAAGCCGAGCTTAAACCTGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1374

QY 61 TCTGACGAGGCTTCTCCGCTCCCGGCCCGCCACTGCTCCCGCTGGCGAGGACGGT 120  
DB 1375 TCTGACGAGGCTTCTCCGCTCCCGGCCCGCCACTGCTCCCGCTGGCGAGGACGGT 1434

QY 121 CGTGTGAGCCCTATGTGACCTTCTGAGTCTTACCGCTCTGAGCGTGAACCATGGC 180  
DB 1435 CGTGTGAGCCCTATGTGACCTTCTGAGTCTTACCGCTCTGAGCGTGAACCATGGC 1494

QY 181 GAGCAGAGCGTGGTGACAGCACCG 204  
DB 1495 GAGCAGAGCGTGGTGACAGCACCG 1518

RESULT 15  
HSU49928  
LOCUS  
DEFINITION Homo sapiens TAK1 binding protein (TAB1) mRNA, complete cds.  
ACCESSION U49928  
VERSION U49928.1 GI:1401125  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 3096)  
Shibuya, H., Yamaguchi, K., Shirakabe, K., Tonegawa, A., Gotoh, Y.,  
Ueno, N., Irie, K., Nishida, E. and Matsumoto, K.  
TAB1: an activator of the TAK1 MAPKK in TGF-beta signal  
transduction  
Science 272 (5265), 1179-1182 (1996)  
96216294  
8638164  
PubMed  
2 (bases 1 to 3096)  
Shibuya, H.  
Direct Submission  
Submitted (25-FEB-1996) Hiroshi Shibuya, Faculty of Pharmaceutical  
Sciences, Hokkaido University, Nishi 6-chome, Kita 12, Kita-ku,  
Sapporo, Hokkaido 060, Japan  
Location/Qualifiers  
1. .3096  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1. .3096  
/gene="TAB1"  
21. .1535  
/gene="TAB1"  
/note="activator for TAK1"  
/codon\_start=1  
/product="TAK1 binding protein"  
/protein\_id="AAC12660.1"  
/db\_xref="GI:1401126"

gene  
CDS  
FEATURES  
source  
BASE COUNT 642 a 936 c 952 g 566 t  
ORIGIN  
ESHPEDSWLKFSENNFLYGVNGVNTNFVAORLSAELLGOLNAEAEADV  
REVLQAQDVVERSELEDDALEAKASLSQSLPGVPQHQLPPOYQKILERLTLER  
EISGAMAVVLLNNKLYVANVGNRALLKSTVDGLQVTLQNVNHTTNEDELFRLL  
SGLDAGIKIQVYRIICGQSTRIGDYKYGYTDIDLLSAKSKPIIAPEFHGAQ  
PLDGVTFGLVMSGLYKALEAAHGFGQANQIAAMIDTEFAKQTSLDAAQAVVDRV  
KRIHSDTFASGGERARFCPRHEDTLLVNFYGLGMSQPTPPAPAGGRVPSV  
PYSSAQSTSKTSVTLVLMPSOGQMVNGAHSATLDEATPTLTNOSPILTLQSTWHT  
QSSSSSDGGLFRSRPLPFEGDGRVEFYVDFAEFRLMSVDHGQSVTAP"

Query Match 100.0%; Score 204; DB 9; Length 3096;  
Best Local Similarity 100.0%; Pred. No. 1.6e-37;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 60  
DB 1329 CAAGCCGAGCTTAAACCTGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1388

QY 61 TCTGACGAGGCTTCTCCGCTCCCGGCCCGCCACTGCTCCCGCTGGCGAGGACGGT 120  
DB 1389 TCTGACGAGGCTTCTCCGCTCCCGGCCCGCCACTGCTCCCGCTGGCGAGGACGGT 1448

QY 121 CGTGTGAGCCCTATGTGACCTTCTGAGTCTTACCGCTCTGAGCGTGAACCATGGC 180  
DB 1449 CGTGTGAGCCCTATGTGACCTTCTGAGTCTTACCGCTCTGAGCGTGAACCATGGC 1508

QY 181 GAGCAGAGCGTGGTGACAGCACCG 204  
DB 1509 GAGCAGAGCGTGGTGACAGCACCG 1532

Search completed: December 4, 2003, 06:11:26  
Job time : 950.378 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 00:43:57 ; Search time 104.986 Seconds  
(without alignments)  
5245.298 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541

Perfect score: 204

Sequence: 1 caaagccgacctaaccct.....agagcgtgtgacagcaccg 204

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	1515	21	Human TAB1 coding
2	204	100.0	1560	18	Human TAB1 (TAK1 b
3	204	100.0	1560	18	Human TAB1 (TAK1 b
4	204	100.0	1560	20	Human TAB1 encodin
5	204	100.0	1560	21	Human TAB1-1 nucleo
6	204	100.0	1568	20	Human TAB1 encodin
7	204	100.0	1569	20	Human TAB1-FLAG en
8	204	100.0	3450	25	Human mddt cDNA SE

9	204	100.0	16877	22	ABA20494	Human nervous syst
10	204	100.0	16877	22	AAL36984	Human musculoskele
11	204	100.0	16877	25	ABX59972	cDNA encoding nove
12	203.6	99.8	696	21	AAF15895	Human prostate can
13	37.2	18.2	670	24	ABO56072	Human ovarian anti
14	37.2	18.2	1485	19	AAV16790	cDNA encoding a hu
15	37.2	18.2	1505	19	AAV39457	Modified human ala
16	37.2	18.2	1648	14	AAQ40502	Human ALT gene. H
17	36.8	18.0	2036	24	ABK27558	DNA encoding Chlam
18	36.4	17.8	591	21	AAA79488	Human cDNA differe
19	36.4	17.8	2380	21	AAA79712	Eucalyptus grandis
20	35.2	17.3	425	22	AAF64749	Novel human polyinu
21	35.2	17.3	31422	21	AAQ2302	S. avermitilis ave
22	35.2	17.3	31422	22	AAH79278	Streptomyces averm
23	35	17.2	5811	25	ABX62899	Human activated T
24	34.2	16.8	148834	24	ABK83570	Human cDNA differe
25	34	16.7	588	25	ABX56048	M. echinospora cal
26	34	16.7	4748	12	AAQ13286	P. denitrificans ge
27	33.8	16.6	36538	25	ABV75558	Saccharopolyspora
28	33.8	16.6	47981	22	AAF30757	Micromonospora meg
29	33.6	16.5	81001	22	AAF30035	Human apolipoprote
30	33.6	16.5	1691080	24	ABX08336	Human phosphodiester
31	33.4	16.4	349980	24	ABQ81849	Bifidobacterium lo
32	33	16.2	460	24	ABK44501	Inhibitory PAS dom
33	33	16.2	623	22	ABA07440	Human pancreatic c
34	33	16.2	623	22	ABA07441	Human pancreatic c
35	33	16.2	623	22	ABA07442	Human pancreatic c
36	33	16.2	623	22	AAK91171	Human digestive sy
37	33	16.2	623	22	AAK91172	Human digestive sy
38	33	16.2	623	22	AAK91173	Human digestive sy
39	33	16.2	1286	21	AAC44066	Zea mays DNA fragm
40	32.8	16.1	2491	22	AAH15154	Human cDNA sequenc
41	32.8	16.1	3911	23	AAS81339	DNA encoding novel
42	32.8	16.1	4229	23	AAS74978	DNA encoding novel
43	32.8	16.1	7282	24	ABN59772	Novel human coding
44	32.8	16.1	8480	20	AAZ27521	Interleukin-2 rece
45	32.4	15.9	14654	22	AAS28620	Genomic sequence #

## ALIGNMENTS

### RESULT 1

AAZ48861  
ID AAZ48861 standard; cDNA; 1515 BP.

XX AAZ48861;

XX AC

XX 24-MAR-2000 (first entry)

XX DE Human TAB1 coding sequence.

XX XX

KW Human; TAB1; XIAP; X-linked inhibitor of apoptosis protein; TGF-beta;  
KW transforming growth factor-beta activated kinase 1; monocyte migration;  
KW TAK1 binding protein 1; extracellular matrix protein production;  
KW cell growth inhibitor; beta-amyloid protein deposition;  
KW immunosuppression; Transforming growth factor-beta; ds.

XX OS Homo sapiens.

XX XX

XX JF11326328-A.

XX PD 26-NOV-1999.

XX PF 13-MAY-1998; 98JP-0130378.

XX PR 13-MAY-1998; 98JP-0130378.

XX PA (MATS/) MATSUMOTO K.

XX XX

XX DR WPI; 2000-078337/07.

XX DR P-PSDB; AAY59450.

XX XX

PT Screening a substance which inhibits combination of the X-linked  
PT inhibitor of apoptosis protein  
XX  
PS Claim 2; Page 25-26; 43pp; Japanese.  
XX  
CC This sequence encodes the human TAB1 protein.  
CC The invention relates to a method for screening a substance inhibiting  
CC the formation of a complex between XIAP and TAB1, in which X-linked  
CC inhibitor of apoptosis protein (XIAP), transforming growth factor-beta  
CC activated kinase 1 (TAK1) binding protein 1 (TAB1) and a substance to be  
CC tested are contacted with each other and then the presence or formation  
CC of a complex between XIAP and TAB1 is detected. The substance can be used  
CC as a drug for extracellular matrix protein production enhancement, cell  
CC growth inhibition, monocytic migration, physiologically active substance  
CC induction, immunosuppression, and beta-amyloid protein deposition. A  
CC substance inhibiting the formation of a complex between TAB1 and XIAP as  
CC well as between XIAP and TGF-beta (Transforming growth factor-beta) type  
CC I and/or type II receptor is useful as a drug.  
XX  
SQ Sequence 1515 BP; 323 A; 457 C; 463 G; 272 T; 0 other;

Query Match 100.0%; Score 204; DB 21; Length 1515;  
Best Local Similarity 100.0%; Pred. No. 3.5e-45;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGCCGACCTTAACCTGCGAGTCACCAACGACGACGAGCAGCTCCAGC 60  
Db 1309 CAAAGCCGACCTTAACCTGCGAGTCACCAACGACGACGAGCAGCTCCAGC 1368  
QY 61 TCTGACGGAGGCTTTCGGCTCCGCGCCGCGCCACTCGCTCCGCTGCGGAGGACGGT 120  
Db 1369 TCTGACGGAGGCTTTCGGCTCCGCGCCGCGCCACTCGCTCCGCTGCGGAGGACGGT 1428  
QY 121 CGTGTGAGCCCTATGTGGACTTTGTGAGTTTACCGCCCTCTGGAGCGTGGACCATGGC 180  
Db 1429 CGTGTGAGCCCTATGTGGACTTTGTGAGTTTACCGCCCTCTGGAGCGTGGACCATGGC 1488  
QY 181 GAGCAGAGCGTGTGACAGCACCG 204  
Db 1489 GAGCAGAGCGTGTGACAGCACCG 1512

RESULT 2  
AAT91175  
ID AAT91175 standard; cDNA; 1560 BP.  
XX AC AAT91175;  
XX AC AAT91175;  
DT 25-MAR-2003 (updated)  
DT 14-APR-1998 (first entry)  
XX DE Human TAB1 (TAK1 binding protein) cDNA.  
XX KW TAB1; TAK1 binding protein; transforming growth factor-beta;  
KW signal transduction; human; ds.  
XX OS Homo sapiens.  
XX PH Key Location/Qualifiers  
XX CD 30..1544  
XX FT /\*tag= a  
XX FT variation 185  
XX FT /\*tag= b  
XX FT /note= "another clone has adenine at position 185,  
XX with codon AGC (Ser) altered to AGA (Arg)"  
XX EN EP803571-A2.  
XX XX 29-OCT-1997.  
XX PD 24-APR-1997; 97EP-0302808.  
XX PF 20-NOV-1996; 96US-0752891.  
XX PR

PR 24-APR-1996; 96JP-0126282.  
PR 28-OCT-1996; 96JP-0300856.  
XX  
XX (CHUS) CHUGAI SEIYAKU KK.  
XX Matsumoto K, Nishida E;  
XX WPI: 1997-515318/48.  
XX P-PSDB; AAN26706.

XX DNA encoding TAK1 binding protein TAB1 - member of transforming  
XX growth factor beta receptor signal production pathway, which  
XX activates TAK-1 kinase activity upon binding  
XX  
XX Claim 1; Page 17-19; 30pp; English.  
XX  
XX This cDNA clone codes for human TAB1 (see AAN26706), a novel member  
XX of the transforming growth factor-beta receptor signal transduction  
XX pathway, which activates TAK-1 kinase activity upon binding. To  
XX obtain the full-length TAB1 coding sequence, a human kidney cDNA  
XX library was screened using as a probe a partial TAB1 cDNA obtained  
XX from a yeast two-hybrid assay for proteins that interacted with  
XX TAK1. The 5' terminus was identified by 5'RACE. 2 Different  
XX clones were sequenced, with cytosine and adenine (see AAT91178) as  
XX the 185th nucleotide, respectively, and deposited as FERM BP-5599  
XX and FERM BP-5508, respectively. Also claimed are: isolated DNA  
XX encoding a protein modified by a substitution, deletion and/or  
XX addition of 1 or more amino acids of the 504-residue TAB1 sequence;  
XX (2) DNA which can hybridize with the 1560 bp nucleic acid sequence;  
XX (3) isolated DNA encoding a protein comprising amino acids 21-579  
XX or 437-504 of the 504 TAB1 sequence; (4) DNA encoding a fusion  
XX protein comprising an above protein or polypeptide; (5) expression  
XX vector comprising an above DNA; and (6) host cell, preferably a  
XX mammalian or yeast cell, transformed by the expression vector.  
XX Cells expressing TAB1 and TAK1 can be used to screen for TGF-beta  
XX signalling pathway inhibitors by contacting the cells with a test  
XX compound, and measuring the TAK1 kinase activity.  
XX (Updated on 25-MAR-2003 to correct PA field.)  
XX  
XX Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Query Match 100.0%; Score 204; DB 18; Length 1560;  
Best Local Similarity 100.0%; Pred. No. 3.5e-45;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGCCGACCTTAACCTGCGAGTCACCAACGACGACGAGCAGCTCCAGC 60  
Db 1338 CAAAGCCGACCTTAACCTGCGAGTCACCAACGACGACGAGCAGCTCCAGC 1397  
QY 61 TCTGACGGAGGCTTTCGGCTCCGCGCCGCGCCACTCGCTCCGCTGCGGAGGACGGT 120  
Db 1398 TCTGACGGAGGCTTTCGGCTCCGCGCCGCGCCACTCGCTCCGCTGCGGAGGACGGT 1457  
QY 121 CGTGTGAGCCCTATGTGGACTTTGTGAGTTTACCGCCCTCTGGAGCGTGGACCATGGC 180  
Db 1458 CGTGTGAGCCCTATGTGGACTTTGTGAGTTTACCGCCCTCTGGAGCGTGGACCATGGC 1517  
QY 181 GAGCAGAGCGTGTGACAGCACCG 204  
Db 1518 GAGCAGAGCGTGTGACAGCACCG 1541

RESULT 3  
AAT91178  
ID AAT91178 standard; cDNA; 1560 BP.  
XX AC AAT91178;  
XX AC AAT91178;  
XX DT 25-MAR-2003 (updated)  
XX DT 14-APR-1998 (first entry)  
XX DE Human TAB1 (TAK1 binding protein) cDNA.  
XX

KW TAB1; TAK1 binding protein; transforming growth factor-beta;  
 XX signal transduction; human; ds.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 30..1544  
 FT /\*tag= a  
 FT variation 185  
 FT /\*tag= b  
 FT /note= "another clone has cytosine at position 185,  
 FT with codon AGA (Arg) altered to AGC (Ser)."  
 XX  
 PN EP803571-A2.  
 XX  
 XX 29-OCT-1997.  
 XX  
 PF 24-APR-1997; 97EP-0302808.  
 XX  
 PR 20-NOV-1996; 96US-0752891.  
 PR 24-APR-1996; 96JP-0126282.  
 PR 28-OCT-1996; 96JP-0300856.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Matsumoto K, Nishida E;  
 XX  
 DR WPI; 1997-515318/48.  
 DR P-PSDB; AAW26707.  
 XX  
 XX  
 FT DNA encoding TAK1 binding protein TAB1 - member of transforming  
 FT growth factor beta receptor signal production pathway, which  
 FT activates TAK-1 kinase activity upon binding  
 FT  
 PS Example 5; Page 19-21; 30pp; English.  
 XX  
 CC This cDNA clone codes for human TAB1 (see AAW26707), a novel member  
 CC of the transforming growth factor-beta receptor signal transduction  
 CC pathway, which activates TAK-1 kinase activity upon binding. To  
 CC obtain the full-length TAB1 coding sequence, a human kidney cDNA  
 CC library was screened using as a probe a partial TAB1 cDNA obtained  
 CC from a yeast two-hybrid assay for proteins that interacted with  
 CC TAK1. The 5' terminus was identified by 5'RACE. 2 Different  
 CC clones were sequenced, with cytosine (see AAT91175) and adenine as  
 CC the 185th nucleotide, respectively, and deposited as FERM BP-5599  
 CC and FERM BP-5508, respectively.  
 CC encoding a protein modified by a substitution, deletion and/or  
 CC addition of 1 or more amino acids of the 504-residue TAB1 sequence;  
 CC (2) DNA which can hybridize with the 1560 bp nucleic acid sequence;  
 CC (3) isolated DNA encoding a protein comprising amino acids 21-579  
 CC or 437-504 of the 504 TAB1 sequence; (4) DNA encoding a fusion  
 CC protein comprising an above protein or polypeptide; (5) expression  
 CC vector comprising an above DNA; and (6) host cell, preferably a  
 CC mammalian or yeast cell, transformed by the expression vector.  
 CC Cells expressing TAB1 and TAK1 can be used to screen for TGF-beta  
 CC signalling pathway inhibitors by contacting the cells with a test  
 CC compound, and measuring the TAK1 kinase activity.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 1560 BP; 333 A; 468 C; 480 G; 279 T; 0 other;  
 Query Match 100.0%; Score 204; DB 18; Length 1560;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-45;  
 Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAAGCCGACCTTAACTTCCAGTCCACCAACGACGACGAGAGCAGCTCCAGC 60  
 DB 1338 CAAGCCGACCTTAACTTCCAGTCCACCAACGACGACGAGAGCAGCTCCAGC 1397  
 QY 61 TCTGACGAGGCGCTCTTCCGCTCCCGCGCGCCGCTCGCTCCCGCTCGGAGGACGGT 120  
 DB 1398 TCTGACGAGGCGCTCTTCCGCTCCCGCGCGCCGCTCGCTCCCGCTCGGAGGACGGT 1457

QY 121 CGTGTGAGCCCTATCTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180  
 DB 1458 CGTGTGAGCCCTATCTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1517  
 QY 181 GAGCAGAGCGTGTGACAGCACC 204  
 DB 1518 GAGCAGAGCGTGTGACAGCACC 1541  
 RESULT 4  
 AAX56278  
 ID AAX56278 standard; DNA; 1560 BP.  
 XX  
 AC AAX56278;  
 XX  
 DT 21-JUL-1999 (first entry)  
 XX  
 DE Human TAB1 encoding DNA.  
 XX  
 KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
 KW transforming growth factor beta; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 30..1544  
 FT /\*tag= a  
 XX  
 PN WO9921010-A1.  
 XX  
 PD 29-APR-1999.  
 XX  
 PF 22-OCT-1998; 98WO-JP04796.  
 XX  
 PR 22-OCT-1997; 97JP-0290188.  
 XX  
 PA (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 PI Ohtomo T, Ono K, Tsuchiya M;  
 DR WPI; 1999-312645/26.  
 DR P-PSDB; AAY09541.  
 XX  
 PT Screening for TGF- beta inhibitory substances, which are useful as  
 PT drugs for treatment of diseases relating to its disorder  
 XX  
 PS Example 1; Page 143-147; 195pp; Japanese.  
 XX  
 CC A method has been developed for screening for substances which inhibit  
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method  
 CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 CC and (b) detecting the amount of bound polypeptide, in which the sample  
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming  
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
 CC indications e.g. as TGF-beta signal transduction inhibitors or  
 CC activators, or extracellular matrix protein production enhancement  
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or  
 CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence encodes human TAB1.  
 XX  
 SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;  
 Query Match 100.0%; Score 204; DB 20; Length 1560;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-45;  
 Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAAGCCGACCTTAACTTCCAGTCCACCAACGACGACGAGAGCAGCTCCAGC 60  
 DB 1338 CAAGCCGACCTTAACTTCCAGTCCACCAACGACGACGAGAGCAGCTCCAGC 1397



CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence encodes human TAB1.  
 XX  
 SQ Sequence 1568 BP; 339 A; 472 C; 477 G; 280 T; 0 other;  
 Query Match 100.0%; Score 204; DB 20; Length 1568;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-45;  
 Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAAAGCCCGACCTTAACTTACCTTGAAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 60  
 Db 1346 CAAAGCCCGACCTTAACTTGAAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1405  
 QY 61 TCTGACGGAGGCTCTTCCGCTCCCGCCGCCCACTCGCTCCGCTCGGAGGACGCT 120  
 Db 1406 TCTGACGGAGGCTCTTCCGCTCCCGCCGCCCACTCGCTCCGCTCGGAGGACGCT 1465  
 QY 121 CGTGTGAGCCCTATGTGACCTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180  
 Db 1466 CGTGTGAGCCCTATGTGACCTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1525  
 QY 181 GAGCAGAGCGTGTGACAGCACCG 204  
 Db 1526 GAGCAGAGCGTGTGACAGCACCG 1549

RESULT 7  
 AAX56282  
 ID AAX56282 standard; DNA; 1569 BP.  
 XX  
 AC AAX56282;  
 XX  
 DT 21-JUL-1999 (first entry)  
 XX  
 DE Human TAB1-FLAG encoding DNA.  
 XX  
 KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
 KW transforming growth factor beta; ss.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 7..1560  
 FT /\*tag= a  
 XX  
 PN WO9921010-A1.  
 XX  
 PD 29-APR-1999.  
 XX  
 PF 22-OCT-1998; 98WO-JP04796.  
 XX  
 PR 22-OCT-1997; 97JP-0290188.  
 XX  
 PA (CHUS) CHUGAI SEIYAKU KK.  
 XX  
 PI Ohtomo T, Ono K, Tsuchiya M;  
 XX  
 DR WPI; 1999-312645/26.  
 DR P-PSDB; AAY09546.  
 XX

Screening for TGF- beta inhibitory substances, which are useful as  
 drugs for treatment of diseases relating to its disorder  
 XX  
 PS Example 1; Page 159-163; 195pp; Japanese.  
 XX

CC A method has been developed for screening for substances which inhibit  
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method  
 CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 CC and (b) detecting the amount of bound polypeptide, in which the sample  
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming  
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for

CC indications e.g. as TGF-beta signal transmission inhibitors or  
 CC activators, or extracellular matrix protein production enhancement  
 CC inhibitors or activators, or cell proliferation prevention inhibitors or  
 CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence encodes TAB1-FLAG from an example of  
 CC the present invention.  
 XX

SQ Sequence 1569 BP; 343 A; 466 C; 476 G; 284 T; 0 other;

Query Match 100.0%; Score 204; DB 20; Length 1569;  
 Best Local Similarity 100.0%; Pred. No. 3.5e-45;  
 Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAAAGCCCGACCTTAACTTACCTTGAAGTCCACCAACAGCAGCAGCAGCTCCAGC 60  
 Db 1315 CAAAGCCCGACCTTAACTTGAAGTCCACCAACAGCAGCAGCAGCTCCAGC 1374  
 QY 61 TCTGACGGAGGCTCTTCCGCTCCCGCCGCCCACTCGCTCCGCTCGGAGGACGCT 120  
 Db 1375 TCTGACGGAGGCTCTTCCGCTCCCGCCGCCCACTCGCTCCGCTCGGAGGACGCT 1434  
 QY 121 CGTGTGAGCCCTATGTGACCTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180  
 Db 1435 CGTGTGAGCCCTATGTGACCTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1494  
 QY 181 GAGCAGAGCGTGTGACAGCACCG 204  
 Db 1495 GAGCAGAGCGTGTGACAGCACCG 1518

RESULT 8  
 ABX34568  
 ID ABX34568 standard; cDNA; 3450 BP.  
 XX  
 AC ABX34568;  
 XX  
 DT 13-FEB-2003 (first entry)  
 XX  
 DE Human mddt cDNA SEQ ID 129.  
 XX  
 KW MDDT; human; disease detection and treatment molecule polypeptide;  
 KW anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;  
 KW haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;  
 KW gene therapy; protein replacement therapy; cell proliferative disorder;  
 KW cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; sarcoma;  
 KW anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
 KW Goodpasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;  
 KW psoriasis; hepatitis; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200279449-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-US09944.  
 XX

28-MAR-2001; 2001US-279619P.  
 29-MAR-2001; 2001US-280067P.  
 29-MAR-2001; 2001US-280068P.  
 16-MAY-2001; 2001US-291280P.  
 17-MAY-2001; 2001US-291829P.  
 17-MAY-2001; 2001US-291849P.  
 19-JUN-2001; 2001US-299428P.  
 20-JUN-2001; 2001US-299776P.  
 20-JUN-2001; 2001US-300001P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX



PR 25-SEP-2000; 2000US-0234398.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 29-SEP-2000; 2000US-0236371.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0242221.  
 PR 01-NOV-2000; 2000US-0242221.  
 PR 08-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0244674.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249246.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 01-DEC-2000; 2000US-0251160.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251038.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251899.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.

XX PA (HUMA-) HUMAN GENOME SCI INC.  
 XX PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI; 2001-541565/60.  
 XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating nervous system  
 PT cancers and metastases -  
 XX Disclosure; SEQ ID NO 12825; 1701pp + Sequence Listing; English.  
 XX The invention relates to novel genes (ABAI1004-ABA21534) and proteins  
 CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 16977 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 204; DB 22; Length 16877;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-45;  
 Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CAAAGCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACGACGACG 60  
 Db 15011 CAAAGCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACGACG 15070  
 QY 61 TCTGACGAGGCGCTTTCGCTCCGCGCCGCGCCACTCGCTCCGCTCCGCGAGGACGGT 120  
 Db 15071 TCTGACGAGGCGCTTTCGCTCCGCGCCGCGCCACTCGCTCCGCTCCGCGAGGACGGT 15130  
 QY 121 CGTGTTCAGCCCTATGTGGACTTTGCTGATTTTACCGCTCTGGAGCGTGGACCAATGGC 180  
 Db 15131 CGTGTTCAGCCCTATGTGGACTTTGCTGATTTTACCGCTCTGGAGCGTGGACCAATGGC 15190  
 QY 181 GAGCAGAGCGTGTGACAGCACCG 204  
 Db 15191 GAGCAGAGCGTGTGACAGCACCG 15214  
 RESULT 10  
 AAL36984  
 ID AAL36984 standard; DNA; 16877 BP.  
 XX AAL36984;  
 XX AC AAL36984;  
 XX XX  
 DT 08-JAN-2002 (first entry)  
 XX Human musculoskeletal system related polynucleotide SEQ ID NO 3349.  
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein;  
 KW musculoskeletal system; ds.  
 XX Homo sapiens.  
 OS

XX	WO200155367-A1.	27-SEP-2000; 2000US-0235836.	PR
PN	02-AUG-2001.	29-SEP-2000; 2000US-0236327.	PR
XX		29-SEP-2000; 2000US-0236367.	PR
XX		29-SEP-2000; 2000US-0236368.	PR
PD		29-SEP-2000; 2000US-0236369.	PR
XX		29-SEP-2000; 2000US-0236370.	PR
PF	17-JAN-2001; 2001WO-US01338.	02-OCT-2000; 2000US-0236802.	PR
XX		02-OCT-2000; 2000US-0237037.	PR
XX		02-OCT-2000; 2000US-0237038.	PR
XX		02-OCT-2000; 2000US-0237039.	PR
XX		02-OCT-2000; 2000US-0237040.	PR
XX		13-OCT-2000; 2000US-0239935.	PR
XX		13-OCT-2000; 2000US-0239937.	PR
XX		20-OCT-2000; 2000US-0240960.	PR
XX		20-OCT-2000; 2000US-0241221.	PR
XX		20-OCT-2000; 2000US-0241785.	PR
XX		20-OCT-2000; 2000US-0241786.	PR
XX		20-OCT-2000; 2000US-0241787.	PR
XX		20-OCT-2000; 2000US-0241808.	PR
XX		20-OCT-2000; 2000US-0241809.	PR
XX		20-OCT-2000; 2000US-0241826.	PR
XX		01-NOV-2000; 2000US-0244617.	PR
XX		08-NOV-2000; 2000US-0246474.	PR
XX		08-NOV-2000; 2000US-0246475.	PR
XX		08-NOV-2000; 2000US-0246476.	PR
XX		08-NOV-2000; 2000US-0246477.	PR
XX		08-NOV-2000; 2000US-0246478.	PR
XX		08-NOV-2000; 2000US-0246523.	PR
XX		08-NOV-2000; 2000US-0246524.	PR
XX		08-NOV-2000; 2000US-0246525.	PR
XX		08-NOV-2000; 2000US-0246526.	PR
XX		08-NOV-2000; 2000US-0246527.	PR
XX		08-NOV-2000; 2000US-0246528.	PR
XX		08-NOV-2000; 2000US-0246532.	PR
XX		08-NOV-2000; 2000US-0246609.	PR
XX		08-NOV-2000; 2000US-0246610.	PR
XX		08-NOV-2000; 2000US-0246611.	PR
XX		08-NOV-2000; 2000US-0246613.	PR
XX		17-NOV-2000; 2000US-0249207.	PR
XX		17-NOV-2000; 2000US-0249208.	PR
XX		17-NOV-2000; 2000US-0249209.	PR
XX		17-NOV-2000; 2000US-0249210.	PR
XX		17-NOV-2000; 2000US-0249211.	PR
XX		17-NOV-2000; 2000US-0249212.	PR
XX		17-NOV-2000; 2000US-0249213.	PR
XX		17-NOV-2000; 2000US-0249214.	PR
XX		17-NOV-2000; 2000US-0249215.	PR
XX		17-NOV-2000; 2000US-0249216.	PR
XX		17-NOV-2000; 2000US-0249217.	PR
XX		17-NOV-2000; 2000US-0249218.	PR
XX		17-NOV-2000; 2000US-0249244.	PR
XX		17-NOV-2000; 2000US-0249245.	PR
XX		17-NOV-2000; 2000US-0249264.	PR
XX		17-NOV-2000; 2000US-0249265.	PR
XX		17-NOV-2000; 2000US-0249297.	PR
XX		17-NOV-2000; 2000US-0249299.	PR
XX		17-NOV-2000; 2000US-0249300.	PR
XX		01-DEC-2000; 2000US-0250160.	PR
XX		01-DEC-2000; 2000US-0250391.	PR
XX		05-DEC-2000; 2000US-0251030.	PR
XX		05-DEC-2000; 2000US-0251988.	PR
XX		05-DEC-2000; 2000US-0256719.	PR
XX		06-DEC-2000; 2000US-0251479.	PR
XX		08-DEC-2000; 2000US-0251856.	PR
XX		08-DEC-2000; 2000US-0251869.	PR
XX		08-DEC-2000; 2000US-0251869.	PR
XX		08-DEC-2000; 2000US-0251989.	PR
XX		08-DEC-2000; 2000US-0251990.	PR
XX		11-DEC-2000; 2000US-0254097.	PR
XX		05-JAN-2001; 2001US-0259678.	PR
XX		(HUMA-) HUMAN GENOME SCI INC.	XX
PA			PA
XX			XX









QY 14 TAACTCTGAGTCCACCAACGACACGACGAGCAGCTCCAGCTCTGACGGAGGCC 73  
 Db 1172 TCACGAGCAGGTCTTCAATGAGGCTTCCTGGATCAGCTGCAACCCAGTGCAGGGGCCA 1231  
 QY 74 TCTTCGGCTCCCGGCGCGCCCACTCGCTCCCGCTGGCGAGGACGGTCTGTTGAGCCCT 133  
 Db 1232 TGTACTCTCTCCCGCGGTGACGCTGCCCGCGGCGGTGGAGCGGCTCAGGAGCTGG 1291  
 QY 134 ATGTGACTTTGCTGAGTTTACCGGCTCTGAGCGCTGACCAATGCGGAGCAGAGCGTGG 193  
 Db 1292 GCTTGGCCCCGATATGTTCTTCTGCTGGCTCGCTCTGGAGGAGACCGGCACTGCGGTGG 1351  
 QY 194 TGACAG 199  
 Db 1352 TGCCAG 1357

RESULT 15  
 AAV39457  
 ID AAV39457 standard; cDNA; 1505 BP.

XX AAV39457;  
 XX AC  
 XX DT  
 XX 15-OCT-1998 (first entry)  
 XX DE Modified human alanine aminotransferase encoding cDNA.  
 XX KW Human; alanine aminotransferase; ALT; modified; liver disease;  
 XX KW viral hepatitis; hepatic cirrhosis; gene therapy; ds.  
 XX OS Homo sapiens.  
 XX OS Synthetic.

XX FH Key Location/Qualifiers  
 XX CDS 9..1499  
 XX FT /\*tag= a  
 XX FT /product= "modified alanine aminotransferase"

XX PN WO9830703-A1.  
 XX PD 16-JUL-1998.  
 XX PF 19-SEP-1997; 97WO-JP03339.  
 XX PR 09-JAN-1997; 97JP-0013282.  
 XX PA (ORIY ) ORIENTAL YEAST CO LTD.  
 XX PI Eto A, Funatsu M, Matsuo Y, Nakamura A, Tanaka T;  
 XX PI Tanase S;  
 XX DR WPI; 1998-399148/34.  
 XX DR P-PSDB; AAW62267.  
 XX PT Gene for modified human alanine amino-transferase with improved  
 XX PT activity - and production of enzyme by culture of Escherichia coli  
 XX PT transformed with gene  
 XX PS Claim 1; Page 13-16/1; 37pp; Japanese.

XX CC The present sequence encodes a modified human alanine aminotransferase  
 XX CC (ALT) which has improved activity over the wild-type enzyme, and in  
 XX CC which the following five changes have been made: 4-SER, 5-THR, 7-ASP,  
 XX CC 39-GLY and 222-ALA. ALT is an enzyme which is depleted or absent from  
 XX CC serum in patients with liver diseases such as viral hepatitis and  
 XX CC hepatic cirrhosis. The recombinant modified ALT may be used in treatment  
 XX CC of these diseases, and the DNA (with a suitable vector) as gene therapy.

XX SQ Sequence 1505 BP; 266 A; 466 C; 507 G; 266 T; 0 other;

Query Match 18.2%; Score 37.2; DB 19; Length 1505;  
 Best Local Similarity 50.0%; Pred. No. 1.4;

Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
 QY 14 TAACTCTGAGTCCACCAACGACACGACGAGCAGCTCCAGCTCTGACGGAGGCC 73  
 Db 1183 TCACGAGCAGGTCTTCAATGAGGCTTCCTGGATCAGCTGCAACCCAGTGCAGGGGCCA 1242  
 QY 74 TCTTCGGCTCCCGGCGCGCCCACTCGCTCCCGCTGGCGAGGACGGTCTGTTGAGCCCT 133  
 Db 1243 TGTACTCTCTCCCGCGGTGACGCTGCCCGCGGCGGTGGAGCGGCTCAGGAGCTGG 1302  
 QY 134 ATGTGACTTTGCTGAGTTTACCGGCTCTGAGCGCTGACCAATGCGGAGCAGAGCGTGG 193  
 Db 1303 GCTTGGCCCCGATATGTTCTTCTGCTGGCTCGCTCTGGAGGAGACCGGCACTGCGGTGG 1362  
 QY 194 TGACAG 199  
 Db 1363 TGCCAG 1368

Search completed: December 4, 2003, 05:02:29  
 Job time : 107.986 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 04:36:53 : Search time 25.2703 Seconds  
(without alignments)  
3563.166 Million cell updates/sec

Title: US-09-830-144-3 COPY\_1338\_1541

Perfect score: 204

Sequence: 1 caaagccgacctaaccct.....agacgtggtagacagaccg 204

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*

2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*

3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*

4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*

5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	1560	2	US-08-752-891-1
2	204	100.0	1560	2	US-08-752-891-5
3	204	100.0	1560	2	US-09-144-178-1
4	204	100.0	1560	2	US-09-144-178-5
5	204	100.0	1560	3	US-09-406-854-1
6	204	100.0	1560	3	US-09-406-854-5
7	204	100.0	1560	4	US-09-529-279-1
8	204	100.0	1560	4	US-10-158-895-1
9	204	100.0	1568	4	US-09-529-279-42
10	204	100.0	1568	4	US-10-158-895-42
11	204	100.0	1569	4	US-09-529-279-10
12	204	100.0	1569	4	US-10-158-895-103
13	37.2	18.2	1278	4	US-08-252-991A-4932
14	37.2	18.2	1491	2	US-08-941-647A-4
15	37.2	18.2	1505	4	US-09-142-481-1
16	37.2	18.2	1878	4	US-09-252-991A-4933
17	37.2	18.2	2211	4	US-09-252-991A-4927
18	35.4	17.4	762	4	US-09-252-991A-3481
19	35.4	17.4	1203	4	US-09-252-991A-3468
20	35.4	17.4	1269	4	US-09-252-991A-3470
21	33.8	16.6	47981	4	US-09-679-279-1
22	33.6	16.5	81001	4	US-09-750-580-1
23	31.2	15.3	4403765	3	US-09-103-840A-2
24	31.2	15.3	4411529	3	US-09-103-840A-1
25	31	15.2	1101	4	US-09-152-060-23
26	31	15.2	1173	3	US-08-706-216-5
27	31	15.2	1473	4	US-09-152-060-43

28	31	15.2	4092	2	US-08-469-537A-106	Sequence 106, Appl
29	30.8	15.1	408	4	US-09-252-991A-16497	Sequence 16497, A
30	30.8	15.1	417	4	US-09-252-991A-16104	Sequence 16104, A
31	30.8	15.1	753	4	US-09-252-991A-16000	Sequence 16000, A
32	30.8	15.1	855	4	US-09-252-991A-16209	Sequence 16209, A
33	30.8	15.1	4411529	3	US-09-103-840A-1	Sequence 1, Appli
34	30.6	15.0	540	4	US-09-252-991A-7104	Sequence 7104, Ap
35	30.6	15.0	1041	2	US-08-580-545B-5	Sequence 5, Appli
36	30.6	15.0	1041	3	US-09-262-653A-5	Sequence 5, Appli
37	30.6	15.0	1494	4	US-09-252-991A-7049	Sequence 7049, Ap
38	30.6	15.0	4236	4	US-09-252-991A-7057	Sequence 7057, Ap
39	30.6	15.0	10023	4	US-09-252-991A-6997	Sequence 6997, Ap
40	30.6	15.0	44377	2	US-08-804-227C-7	Sequence 7, Appli
41	30.6	15.0	44377	2	US-08-804-198-1	Sequence 1, Appli
42	30.6	15.0	4403765	3	US-09-103-840A-2	Sequence 2, Appli
43	30.4	14.9	507	4	US-09-252-991A-15900	Sequence 15900, A
44	30.4	14.9	639	4	US-09-252-991A-15846	Sequence 15846, A
45	30.4	14.9	1063	3	US-09-077-675A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-752-891-1  
; Sequence 1, Application US/08752891  
; Patent No. 5837819  
; GENERAL INFORMATION: - ~~10202~~ 11/17/98  
; APPLICANT: MATSUMOTO, Kunihiro  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1560 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 30..1541  
; FEATURE:  
; NAME/KEY: mat\_peptide

NAME/KEY: mat\_peptide  
LOCATION: 30..1541  
US-08-752-891-1

Query Match 100.0%; Score 204; DB 2; Length 1560;  
Best Local Similarity 100.0%; Pred. No. 3e-45;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGACGACGACGACGACG 60  
Db 1338 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGACGACGACGACG 1397

QY 61 TCTGACGAGGCGCTCTTCCGCTCCCGCGCCGCGCCACTCGCTCCGCGCTGCGGAGGACGGT 120  
Db 1398 TCTGACGAGGCGCTCTTCCGCTCCCGCGCCGCGCCACTCGCTCCGCGCTGCGGAGGACGGT 1457

QY 121 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180  
Db 1458 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1517

QY 181 GAGCAGGCGTGGTGACAGCAGC 204  
Db 1518 GAGCAGGCGTGGTGACAGCAGC 1541

RESULT 2  
US-08-752-891-5  
; Sequence 5, Application US/08752891  
; Patent No. 5837819  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1560 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 30..1541  
; FEATURE:

NAME/KEY: mat\_peptide  
LOCATION: 30..1541  
US-08-752-891-5

Query Match 100.0%; Score 204; DB 2; Length 1560;  
Best Local Similarity 100.0%; Pred. No. 3e-45;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGACGACGACGACG 60  
Db 1338 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGACGACGACG 1397

QY 61 TCTGACGAGGCGCTCTTCCGCTCCCGCGCCGCGCCACTCGCTCCGCGCTGCGGAGGACGGT 120  
Db 1398 TCTGACGAGGCGCTCTTCCGCTCCCGCGCCGCGCCACTCGCTCCGCGCTGCGGAGGACGGT 1457

QY 121 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180  
Db 1458 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1517

QY 181 GAGCAGGCGTGGTGACAGCAGC 204  
Db 1518 GAGCAGGCGTGGTGACAGCAGC 1541

RESULT 3  
US-09-144-178-1  
; Sequence 1, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1560 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:

```

; STRANDEDNESS: double
; TOPOLOGY: linear
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..1541
;
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 30..1541
;
US-09-144-178-5
Query Match 100.0%; Score 204; DB 2; Length 1560;
Best Local Similarity 100.0%; Pred. No. 3e-45;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGCCCGACCTTAACTCGTGCAGTCCACCAACAGCACGACGAGCAGCAGCTCCAGC 60
Db 1338 CAAAGCCCGACCTTAACTCGTGCAGTCCACCAACAGCACGACGAGCAGCAGCTCCAGC 1397

QY 61 TCTGACGGAGGCTCTTCGGCTCCCGGCCGCCACTCGCTCCCGCTCCGCGAGGACGGT 120
Db 1398 TCTGACGGAGGCTCTTCGGCTCCCGGCCGCCACTCGCTCCCGCTCCGCGAGGACGGT 1457

QY 121 CGTGTGTGACCCCTATGTGACATTTGCTGAGTTTACCGCTCTGGAGCGTGGACCAATGCC 180
Db 1458 CGTGTGTGACCCCTATGTGACATTTGCTGAGTTTACCGCTCTGGAGCGTGGACCAATGCC 1517

QY 181 GAGCAGAGCGTGGTGACGACCG 204
Db 1518 GAGCAGAGCGTGGTGACGACCG 1541

```

RESULT 5  
 US-09-406-854-1  
 ; Sequence 1, Application US/09406854  
 ; Patent No. 6140042  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MATSUMOTO, Kunihiko  
 ; APPLICANT: NISHIDA, Eisuke  
 ; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 3000 K Street, N.W., Suite 500  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/406,854  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/752,891  
 ; FILING DATE: 20-NOV-1996  
 ; APPLICATION NUMBER: JP 8-300856  
 ; FILING DATE: 28-OCT-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 8-126282  
 ; FILING DATE: 24-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 17981/111  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 672-5300  
 ; TELEFAX: (202) 672-5399  
 ; TELEX: 904136  
 ; INFORMATION FOR SEQ ID NO: 1:

```

; NAME/KEY: CDS
; LOCATION: 30...1541
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 30...1541
US-09-144-178-1

Query Match          100.0%; Score 204; DB 2; Length 1560;
Best Local Similarity 100.0%; Pred. No. 3e-45;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  CAAAGCCCGACCTTAACCTCGAGTCCACCAACACGACACGACGAGCAGCAGCTCCAGC 60
Db      1338 CAAAGCCCGACCTTAACCTCGAGTCCACCAACACGACACGACGAGCAGCAGCTCCAGC 1397

QY      61  TCTGACGGAGGCGCTCTTCGGCTCCCGGCCGCCCACTCGCTCCCGCCTGGCGAGGACGGT 120
Db      1398 TCTGACGGAGGCGCTCTTCGGCTCCCGGCCGCCCACTCGCTCCCGCCTGGCGAGGACGGT 1457

QY      121 CGTGTGTAGCCCTATGTGCACTTTGTGTAGTTTTACCGCTCTTGAGCGCTGTGACCAATGCC 180
Db      1458 CGTGTGTAGCCCTATGTGCACTTTGTGTAGTTTTACCGCTCTTGAGCGCTGTGACCAATGCC 1517

QY      181 GAGCAGAGCGTGGTGACAGCACC 204
Db      1518 GAGCAGAGCGTGGTGACAGCACC 1541

```

RESULT 4  
US-09-144-178-5  
; Sequence 5, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/144,178  
; APPLICATION NUMBER: US/09/144,178  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/752,891  
; FILING DATE: 20-NOV-1996  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1560 base pairs  
; TYPE: nucleic acid

95

102e  
Sue Parent

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..1541
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 30..1541
;
US-09-406-854-1

Query Match          100.0%; Score 204; DB 3; Length 1560;
Best Local Similarity 100.0%; Pred. No. 3e-45;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACGACG 60
Db 1338 CAAAGCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 1397

QY 61 TCTGACGGAGCGCTTTCCTCCGCTCCCGCCGCGCCACTCGCTCCCGCTCGGAGGACGGT 120
Db 1398 TCTGACGGAGCGCTTTCCTCCGCTCCCGCCGCGCCACTCGCTCCCGCTCGGAGGACGGT 1457

QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1517

QY 181 GAGCAGAGCGTGTGACAGCAGCG 204
Db 1518 GAGCAGAGCGTGTGACAGCAGCG 1541
```

```

RESULT 6
US-09-406-854-5
; Sequence 5, Application US/09406854
; Patent No. 6140042
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/406.854
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/752.891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
```

```

; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..1541
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 30..1541
;
US-09-406-854-5

Query Match          100.0%; Score 204; DB 3; Length 1560;
Best Local Similarity 100.0%; Pred. No. 3e-45;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 60
Db 1338 CAAAGCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 1397

QY 61 TCTGACGGAGCGCTTTCCTCCGCTCCCGCCGCGCCACTCGCTCCCGCTCGGAGGACGGT 120
Db 1398 TCTGACGGAGCGCTTTCCTCCGCTCCCGCCGCGCCACTCGCTCCCGCTCGGAGGACGGT 1457

QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 180
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 1517

QY 181 GAGCAGAGCGTGTGACAGCAGCG 204
Db 1518 GAGCAGAGCGTGTGACAGCAGCG 1541
```

```

RESULT 7
US-09-529-279-1
; Sequence 1, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JF98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
;
US-09-529-279-1

Query Match          100.0%; Score 204; DB 4; Length 1560;
Best Local Similarity 100.0%; Pred. No. 3e-45;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 60
Db 1338 CAAAGCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 1397

QY 61 TCTGACGGAGCGCTTTCCTCCGCTCCCGCCGCGCCACTCGCTCCCGCTCGGAGGACGGT 120
```



1398 TCTGACGAGGCTCTTCCGCTCCCGCCGACCTCGCTCCCGCTGGCGAGGACGGT 1457  
121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180  
1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517  
181 GAGCAGAGCGTGTGACAGCACC 204  
1518 GAGCAGAGCGTGTGACAGCACC 1541

Query Match 100.0%; Score 204; DB 4; Length 1568;  
Best Local Similarity 100.0%; Pred. No. 3e-45; Indels 0; Gaps 0;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 60  
1346 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 1405

61 TCTGACGAGGCTCTTCCGCTCCCGCCGACCTCGCTCCCGCTGGCGAGGACGGT 120  
1406 TCTGACGAGGCTCTTCCGCTCCCGCCGACCTCGCTCCCGCTGGCGAGGACGGT 1465

121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180  
1466 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1525

181 GAGCAGAGCGTGTGACAGCACC 204  
1526 GAGCAGAGCGTGTGACAGCACC 1549

RESULT 10  
US-10-158-895-42  
; Sequence 42, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278 US/10/158,895  
; CURRENT APPLICATION NUMBER: 2002-06-03  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 42  
; LENGTH: 1568  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (11)..(1549)  
US-10-158-895-42

Query Match 100.0%; Score 204; DB 4; Length 1568;  
Best Local Similarity 100.0%; Pred. No. 3e-45;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 60  
1346 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 1405

1398 TCTGACGAGGCTCTTCCGCTCCCGCCGACCTCGCTCCCGCTGGCGAGGACGGT 1457  
121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180  
1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517  
181 GAGCAGAGCGTGTGACAGCACC 204  
1518 GAGCAGAGCGTGTGACAGCACC 1541

Query Match 100.0%; Score 204; DB 4; Length 1560;  
Best Local Similarity 100.0%; Pred. No. 3e-45;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGAGCAGCTCCAGC 60  
1338 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGAGCAGCTCCAGC 1397

61 TCTGACGAGGCTCTTCCGCTCCCGCCGACCTCGCTCCCGCTGGCGAGGACGGT 120  
1398 TCTGACGAGGCTCTTCCGCTCCCGCCGACCTCGCTCCCGCTGGCGAGGACGGT 1457

121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180  
1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517

181 GAGCAGAGCGTGTGACAGCACC 204  
1518 GAGCAGAGCGTGTGACAGCACC 1541

RESULT 9  
US-09-529-279-42  
; Sequence 42, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279

Query Match 100.0%; Score 204; DB 4; Length 1560;  
Best Local Similarity 100.0%; Pred. No. 3e-45;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGAGCAGCTCCAGC 60  
1338 CAAAGCCGACCTTAACCTGAGTCCACCAACGACGACGAGCAGCTCCAGC 1397

61 TCTGACGAGGCTCTTCCGCTCCCGCCGACCTCGCTCCCGCTGGCGAGGACGGT 120  
1398 TCTGACGAGGCTCTTCCGCTCCCGCCGACCTCGCTCCCGCTGGCGAGGACGGT 1457

121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180  
1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517

181 GAGCAGAGCGTGTGACAGCACC 204  
1518 GAGCAGAGCGTGTGACAGCACC 1541

Qy	61	TCTGACGAGGACCTCTTTCGCTCCGCGCCGCCCACTCGCTCCGCGCTTGGCGAGACGGT	120
Db	1406	TC <sup>1</sup> TGACGAGGACCTCTTTCGCTCCGCGCCGCCCACTCGCTCCGCGCTTGGCGAGACGGT	1455
Qy	121	CGTGTTGAGCCCTATGTGGACTTTTGTCTGAGTTTACCGCCCTCTCGAGCGTGGACCATGGC	180
Db	1466	CGTGTTGAGCCCTATGTGGACTTTTGTCTGAGTTTACCGCCCTCTCGAGCGTGGACCATGGC	1525
Qy	181	GAGCAGAGCGTGTGACAGCACCG	204
Db	1526	GAGCAGAGCGTGTGACAGCACCG	1549

## RESULT 11

```

US-09-529-279-10
; Sequence 10, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHYOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1557)
; US-09-529-279-10

```

Query Match	100.0%;	Score 204;	DB 4;	Length 1569;
Rest Local Similarity	100.0%;	Pred. No. 3e-45;		
Matches 204;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	CAAGCCCGACCTTAA	CCCTGCAGTCCACCA	CACGCA	CACGACGAGCAGGTCCAGC 60
1315	CAAGCCCGACCTTAA	CCCTGCAGTCCACCA	CACGCA	CACGACGAGCAGGTCCAGC 1374
61	CTTGACGGAGCGCTCT	TCGCGTCCCGGGCCGCGCC	CACTCGCTCCCGCCTGGCGAGCAGGT 120	
1375	CTTGACGGAGCGCTCT	TCGCGTCCCGGGCCGCGCC	CACTCGCTTCGCGAGCAGCGT 1434	
121	CGTGTTCGAGCCCTAT	GTGACATTCCTGAGTTTAC	CGCCCTCTGGAGCGTGGACCATGGC 180	
1435	CGTGTTCGAGCCCTAT	GTGACATTCCTGAGTTTAC	CGCCCTCTGGAGCGTGGACCATGGC 1494	
181	GAGCAGAGCGTGGT	GACAGCACC	204	
1495	GAGCAGAGCGTGGT	GACAGCACC	1518	

RESULT 12

US-10-158-895-10  
; Sequence 10, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHITAKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895

```

; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/230188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1557)
US-10-158-895-10

Query Match
Best Local Similarity 100.0%; Score 204; DB 4; Length 1569;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCCTTAACCTCGACTCCACCAACGACGACGAGCAGCAGCTCCAGC 60
Db 1315 CAAAGCCGAGCCTTAACCTGCACTCCACCAACGACGACGAGCAGCAGCTCCAGC 1374
QY 61 TCTGACGAGGACCTCTTCCGCTCCGCGCCGCGCCCACTCGTTCGCGCTGGCAGGACGGT 120
Db 1375 TCTGACGAGGACCTCTTCCGCTCCGCGCCGCGCCCACTCGTTCGCGCTGGCAGGACGGT 1434
QY 121 CGTGTTCAGCCCTATGTGGAGCTTTGCTGAGCTTTTACGCGCTCTGGAGCGTGGACCATGCC 180
Db 1435 CGTGTTCAGCCCTATGTGGAGCTTTGCTGAGCTTTTACGCGCTCTGGAGCGTGGACCATGCC 1494
QY 181 GAGCAGAGCGTGGTGACGACCG 204
Db 1495 GAGCAGAGCGTGGTGACGACCG 1518

```

## RESULT 13

US-09-252-991A-4932/c  
; Sequence 4932, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 4932  
; LENGTH: 1278  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-4932

	Query Match	18.2%;	Score 37.2;	DB 4;	Length 1278;
	Best Local Similarity	57.9%;	Pred. No. 0.29;		
	Matches 66;	Conservative 0;	Mismatches 48;	Indels 0;	Gaps 0;
Qy	64	GACGAGGCGCTCTTCGCTCCGCGCCGCGCCCACTCGCTCCGCTCCGCGAGGACGGTCGT	123		
Db	691	GACGAACGCGCTCTTCGCGCGCGCGCGCGCAGGTGCACGGCCCTTGGCGAAGGCCTCCTT	632		
Qy	124	GTGTAGCCCTATGTGACATTTCTGAGTTTACCGCCTCTGGAGCGTGGACCAT	177		
Db	631	GCCGAAACGGTCTCGCCGTTGAGCAGCAGCGGCAGCTCGCGCGGTAGACCTT	578		

RESULT 15  
US-09-142-481-1  
; Sequence 1, Application US/09142481  
; Patent No. 6316238  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, Atsuo  
; APPLICANT: TANAKA, Toshio  
; APPLICANT: MATSUO, Yushi  
; APPLICANT: TANASE, Sumio  
; APPLICANT: FUNATSU, Masahiko  
; APPLICANT: ETO, Akira  
; TITLE OF INVENTION: PROCESS FOR PRODUCING ACTIVE HUMAN ALT  
; FILE REFERENCE: NAKAMURA=14  
; CURRENT FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: PCT/JP97/03339  
; PRIOR FILING DATE: 1997-09-19  
; PRIOR APPLICATION NUMBER: JP 9-13282  
; PRIOR FILING DATE: 1997-01-09  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1505  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (9)..(1496)  
US-09-142-481-1

Query Match 18.2%; Score 37.2; DB 4; Length 1505;  
Best Local Similarity 50.0%; Pred. No. 0.3;  
Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
QY 14 TAAACCTGAGTCCACCAACGACGACGAGCAGCTCCAGCTCTGAGCGGCG 73  
Db 1183 TCACCGAGCAGCTCTTCAATGAGGCTCTGCGCATCAGCTGCAACCCAGTGCAGGGCGCCA 1242  
QY 74 TCTTCGGCTCCCGCGCGCCACTCGCTCCCGCTGCGGAGGACGCTCGTTGAGCCCT 133  
Db 1243 TGTACTCTCTCCCGCGCGTGCAGCTGCCCGCGCGGTGGAGCGGCTCAGGAGCTGG 1302  
QY 134 ATGTGAGCTTTGCTGAGTTTACCGCTCTGAGGCGTGGACCATGCGGAGCAGAGCGTGG 193  
Db 1303 GCCTGGCCCCCGATATGTTCTCTGCTGCGCTCTCTGGAGGAGACCGGCATCTGCGTGG 1362  
QY 194 TGACAG 199  
Db 1363 TGCCAG 1368  
Search completed: December 4, 2003, 07:06:45  
Job time : 34.2703 secs

RESULT 14  
US-08-941-647A-4  
; Sequence 4, Application US/08941647A  
; Patent No. 5952211  
; GENERAL INFORMATION:  
; APPLICANT: NAKAMURA, ATSUO  
; APPLICANT: TANAKA, TOSHIO  
; APPLICANT: MATSUO, YUSHI  
; APPLICANT: TANASE, SUMIO  
; APPLICANT: FUNATSU, MASAHIKO  
; APPLICANT: ETO, AKIRA  
; TITLE OF INVENTION: METHOD FOR PRODUCING ACTIVE HUMAN  
; TITLE OF INVENTION: ALANINE AMINOTRANSFERASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/941,647A  
; FILING DATE: 30-SEP-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/533,717  
; FILING DATE: 26-SEP-1995  
; APPLICATION NUMBER: JP 6-269119  
; FILING DATE: 07-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 7335-001-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1491 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-941-647A-4

Query Match 18.2%; Score 37.2; DB 2; Length 1491;  
Best Local Similarity 50.0%; Pred. No. 0.3;  
Matches 93; Conservative 0; Mismatches 93; Indels 0; Gaps 0;  
QY 14 TAAACCTGAGTCCACCAACGACGACGAGCAGCTCCAGCTCTGAGCGGCG 73  
Db 1175 TCACCGAGCAGCTCTTCAATGAGGCTCTGCGCATCAGCTGCAACCCAGTGCAGGGCGCCA 1234  
QY 74 TCTTCGGCTCCCGCGCGCCACTCGCTCCCGCTGCGGAGGACGCTCGTTGAGCCCT 133  
Db 1235 TGTACTCTCTCCCGCGGTGAGCTGCCCGCGCGGTGGAGCGGCTCAGGAGCTGG 1294  
QY 134 ATGTGAGCTTTGCTGAGTTTACCGCTCTGAGGCGTGGACCATGCGGAGCAGAGCGTGG 193  
Db 1295 GCCTGGCCCCCGATATGTTCTCTGCTGCGCTCTCTGGAGGAGACCGGCATCTGCGTGG 1354  
QY 194 TGACAG 199  
Db 1355 TGCCAG 1360

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: December 4, 2003, 06:12:04 ; Search time 304.392 Seconds  
(without alignments)  
2227.439 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541  
Perfect score: 204  
Sequence: 1 caaagccgcacctaaccct.....agagcgtggtagacagcagg 204

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues 4403344

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq\*
- 16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	204	100.0	1560	12	US-10-384-743-1
2	204	100.0	1560	13	Sequence 1, Appli
3	204	100.0	1560	13	Sequence 1, Appli
4	204	100.0	1560	13	Sequence 5, Appli
5	204	100.0	1568	12	Sequence 1, Appli
6	204	100.0	1568	13	Sequence 42, Appli
7	204	100.0	1569	13	Sequence 10, Appl
8	204	100.0	1569	13	Sequence 10, Appl
9	204	100.0	16877	10	Sequence 3349, Ap
10	203.6	99.8	696	10	Sequence 330, App
11	199.2	97.6	409	11	Sequence 32946, A
12	36.8	18.0	2036	10	Sequence 17, Appl
13	36.4	17.8	591	14	Sequence 289, App
14	36.4	17.8	2380	14	Sequence 864, App
15	36.2	17.7	2922	14	Sequence 1749, Ap
16	36.2	17.7	9025608	14	Sequence 1, Appli

c 17	35.4	17.4	975	12	US-10-027-632-10526	Sequence 10526, A
c 18	35.4	17.4	975	13	US-10-027-632-10526	Sequence 10526, A
c 19	35.2	17.3	14643	14	US-10-156-761-931	Sequence 931, App
c 20	35	17.2	5811	13	US-10-002-600-15	Sequence 15, Appl
c 21	34.8	17.1	1275	14	US-10-156-761-5872	Sequence 5872, Ap
c 22	34.4	16.9	1089	14	US-10-156-761-6553	Sequence 6553, Ap
c 23	34.4	16.9	1287	10	US-09-887-576-806	Sequence 806, Appli
c 24	34.4	16.9	9025608	14	US-10-156-761-1	Sequence 1, Appli
c 25	33.8	16.6	1011	12	US-10-259-165-521	Sequence 521, App
c 26	33.6	16.5	1128	14	US-10-156-761-6376	Sequence 6376, Ap
c 27	33.6	16.5	547	12	US-10-029-386-12487	Sequence 12487, A
c 28	33.6	16.5	933	14	US-10-156-761-6779	Sequence 6779, Ap
c 29	33.6	16.5	81001	10	US-09-751-877-1	Sequence 1, Appli
c 30	33.6	16.5	81001	11	US-09-842-364-1	Sequence 1, Appli
c 31	33.6	16.5	1691139	14	US-10-067-514-1	Sequence 1, Appli
c 32	33	16.2	460	10	US-09-896-791B-1	Sequence 1, Appli
c 33	33	16.2	28066	12	US-10-017-161-2395	Sequence 2395, Ap
c 34	32.8	16.1	1125	14	US-10-156-761-6007	Sequence 6007, Ap
c 35	32.4	15.9	2640	12	US-09-814-353-19681	Sequence 19681, A
c 36	32.4	15.9	14654	9	US-09-764-860-1054	Sequence 1054, Ap
c 37	32.4	15.9	14654	12	US-10-212-872-1054	Sequence 1054, Ap
c 38	32.4	15.9	14654	14	US-10-074-095-1054	Sequence 1054, Ap
c 39	32.2	15.8	1737	14	US-10-156-761-3579	Sequence 3579, Ap
c 40	31.8	15.6	566	10	US-09-917-800A-960	Sequence 960, App
c 41	31.8	15.6	693	14	US-10-156-761-6496	Sequence 6496, Ap
c 42	31.6	15.5	2838	14	US-10-156-761-6936	Sequence 6936, Ap
c 43	31.4	15.4	1740	14	US-10-156-761-1087	Sequence 1087, Ap
c 44	31.2	15.3	858	12	US-10-027-632-28569	Sequence 28569, A
c 45	31.2	15.3	858	13	US-10-027-632-28569	Sequence 28569, A

ALIGNMENTS

RESULT 1

- US-10-384-743-1
- Sequence 1, Application US/10384743
- Publication No. US20030162228A1
- GENERAL INFORMATION:
- APPLICANT: ONO, KOICHIRO
- APPLICANT: OHTOMO, TOSHIHIKO
- APPLICANT: TSUCHIYA, MAGAYUKI
- TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
- FILE REFERENCE: 053466/0278
- CURRENT APPLICATION NUMBER: US/10/384,743
- CURRENT FILING DATE: 2003-03-11
- PRIOR APPLICATION NUMBER: US/09/529,279
- PRIOR FILING DATE: 2000-04-11
- PRIOR APPLICATION NUMBER: PCT/JP98/04796
- PRIOR FILING DATE: 1998-10-22
- PRIOR APPLICATION NUMBER: JP 9/290188
- PRIOR FILING DATE: 1997-10-22
- NUMBER OF SEQ ID NOS: 48
- SOFTWARE: PatentIn Ver. 2.1
- SEQ ID NO 1
- LENGTH: 1560
- TYPE: DNA
- ORGANISM: Homo sapiens
- FEATURE:
- NAME/KEY: CDS
- LOCATION: (30)...(1541)

US-10-384-743-1

Query Match	100.0%	Score	204	DB	12	Length	1560
Best Local Similarity	100.0%	Pred. No.	1.7e-51				
Matches	204	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	CAAAGCCGACCTTAACCTTCAGTCCACCAACGACGACGACGAGCTCCAGC	60				
DB	1338	CAAAGCCGACCTTAACCTTCAGTCCACCAACGACGACGACGAGCTCCAGC	1397				
QY	61	TCTGACGAGGCGCTTCTCGCTCCGCGGCCGCCACTCGCTCCCGGAGGACCGT	120				

Db 1398 TCTGACGAGCGCTCTTCCGCTCCGCGCCGCGCCACTCGCTCCGCTGGCGGAGGACGGT 1457  
QY 121 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTCTGGAGCGTGAGCATTGGC 180  
Db 1458 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTCTGGAGCGTGAGCATTGGC 1517  
QY 181 GAGCAGAGCGTGTGTGACAGCACCG 204  
Db 1518 GAGCAGAGCGTGTGTGACAGCACCG 1541

## RESULT 2

US-10-123-427-1  
; Sequence 1, Application US/10123427  
; Publication No. US20020119525A1  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/123,427  
FILING DATE: 17-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/406,854  
FILING DATE: <Unknown>  
FILING DATE: 20-NOV-1996  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1560 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 30..1541  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 30..1541

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-123-427-1

Query Match 100.0%; Score 204; DB 13; Length 1560;  
Best Local Similarity 100.0%; Pred. No. 1.7e-51;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAAGCCGACCTTAACCTGAGTCCACACGACGACGAGCAGCAGCTCCAGC 60

Db 1338 CAAAGCCGACCTTAACCTGAGTCCACACGACGACGAGCAGCAGCTCCAGC 1397  
QY 61 TCTGACGAGCGCTCTTCCGCTCCGCGCCGCGCCACTCGCTCCGCTGGCGGAGGACGGT 120  
Db 1398 TCTGACGAGCGCTCTTCCGCTCCGCGCCGCGCCACTCGCTCCGCTGGCGGAGGACGGT 1457  
QY 121 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTCTGGAGCGTGAGCATTGGC 180  
Db 1458 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTCTGGAGCGTGAGCATTGGC 1517  
QY 181 GAGCAGAGCGTGTGTGACAGCACCG 204  
Db 1518 GAGCAGAGCGTGTGTGACAGCACCG 1541

## RESULT 3

US-10-123-427-5  
; Sequence 5, Application US/10123427  
; Publication No. US20020119525A1  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/123,427  
FILING DATE: 17-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/406,854  
FILING DATE: <Unknown>  
FILING DATE: 20-NOV-1996  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136

INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1560 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 30..1541

FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 30..1541  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-10-123-427-5

```
Query Match      100.0%; Score 204; DB 13; Length 1560;
Best Local Similarity 100.0%; Pred. No. 1.7e-51;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 60
Db 1338 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 1397

QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 120
Db 1398 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 1457

QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517

QY 181 GAGCAGAGCGTGGTGACAGCACCG 204
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 4
US-10-158-895-1
; Sequence 42, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
US-10-158-895-1

Query Match      100.0%; Score 204; DB 13; Length 1560;
Best Local Similarity 100.0%; Pred. No. 1.7e-51;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 60
Db 1338 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 1397

QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 120
Db 1398 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 1457

QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517

QY 181 GAGCAGAGCGTGGTGACAGCACCG 204
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 5
US-10-384-743-42
; Sequence 42, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
US-10-158-895-1

Query Match      100.0%; Score 204; DB 13; Length 1560;
Best Local Similarity 100.0%; Pred. No. 1.7e-51;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 60
Db 1338 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 1397

QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 120
Db 1398 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 1457

QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517

QY 181 GAGCAGAGCGTGGTGACAGCACCG 204
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 6
US-10-158-895-42
; Sequence 42, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
US-10-158-895-42

Query Match      100.0%; Score 204; DB 13; Length 1560;
Best Local Similarity 100.0%; Pred. No. 1.7e-51;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 60
Db 1338 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 1397

QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 120
Db 1398 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 1457

QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517

QY 181 GAGCAGAGCGTGGTGACAGCACCG 204
Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 7
US-10-384-743-42
; Sequence 42, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(1549)
US-10-384-743-42

Query Match      100.0%; Score 204; DB 12; Length 1568;
Best Local Similarity 100.0%; Pred. No. 1.7e-51;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 60
Db 1346 CAAGCCGAGCTTAAACCTGAGTCCACCAACGACGACGACGAGCAGTCCAGC 1405

QY 61 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 120
Db 1406 TCTGACGAGGCGCTTCCGCTCCGCGCCGCGCCACTCGCTCCCGCTGGCGAGACCGT 1465

QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 180
Db 1466 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1525

QY 181 GAGCAGAGCGTGGTGACAGCACCG 204
Db 1526 GAGCAGAGCGTGGTGACAGCACCG 1549

RESULT 8
US-10-158-895-42
; Sequence 42, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(1549)
US-10-158-895-42
```

; NAME/KEY: CDS  
; LOCATION: (11)..(1549)  
US-10-158-895-42

Query Match 100.0%; Score 204; DB 13; Length 1569;  
Best Local Similarity 100.0%; Pred. No. 1.7e-51;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGAGAGAGTCCAGC 60  
Db 1346 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGAGAGAGTCCAGC 1405  
  
QY 61 TCTGACGAGGCGCTTCCGCTCCCGCCGCGCCCACTCGCTCCCGCTGGGAGGAGCGT 120  
Db 1406 TCTGACGAGGCGCTTCCGCTCCCGCCGCGCCCACTCGCTCCCGCTGGGAGGAGCGT 1465  
  
QY 121 CGTGTGAGCCCTAATGAGTCTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 180  
Db 1466 CGTGTGAGCCCTAATGAGTCTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 1525  
  
QY 181 GAGCAGAGCGTGGTGACGACCG 204  
Db 1526 GAGCAGAGCGTGGTGACGACCG 1549

## RESULT 7

US-10-384-743-10  
; Sequence 10, Application US/10384743  
; Publication No. US20030162228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 1569  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1557)  
US-10-384-743-10

Query Match 100.0%; Score 204; DB 12; Length 1569;  
Best Local Similarity 100.0%; Pred. No. 1.7e-51;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGAGAGTCCAGC 60  
Db 1315 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGAGAGTCCAGC 1374  
  
QY 61 TCTGACGAGGCGCTTCCGCTCCCGCCGCGCCCACTCGCTCCCGCTGGGAGGAGCGT 120  
Db 1375 TCTGACGAGGCGCTTCCGCTCCCGCCGCGCCCACTCGCTCCCGCTGGGAGGAGCGT 1434  
  
QY 121 CGTGTGAGCCCTAATGAGTCTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 180  
Db 1435 CGTGTGAGCCCTAATGAGTCTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 1494  
  
QY 181 GAGCAGAGCGTGGTGACGACCG 204  
Db 1495 GAGCAGAGCGTGGTGACGACCG 1518

## RESULT 8

US-10-158-895-10  
; Sequence 10, Application US/10158895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 1569  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1557)  
US-10-158-895-10

Query Match 100.0%; Score 204; DB 13; Length 1569;  
Best Local Similarity 100.0%; Pred. No. 1.7e-51;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGAGAGTCCAGC 60  
Db 1315 CAAAGCCGACCTTAACCTGAGTCCAGTCCACCAACGACGACGAGAGAGTCCAGC 1374  
  
QY 61 TCTGACGAGGCGCTTCCGCTCCCGCCGCGCCCACTCGCTCCCGCTGGGAGGAGCGT 120  
Db 1375 TCTGACGAGGCGCTTCCGCTCCCGCCGCGCCCACTCGCTCCCGCTGGGAGGAGCGT 1434  
  
QY 121 CGTGTGAGCCCTAATGAGTCTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 180  
Db 1435 CGTGTGAGCCCTAATGAGTCTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 1494  
  
QY 181 GAGCAGAGCGTGGTGACGACCG 204  
Db 1495 GAGCAGAGCGTGGTGACGACCG 1518

## RESULT 9

US-09-764-877-3349  
; Sequence 3349, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3349  
; LENGTH: 16877  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-3349  
  
Query Match 100.0%; Score 204; DB 10; Length 16877;  
Best Local Similarity 100.0%; Pred. No. 1.9e-51;



	Matches	204;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	CAAAGCCGACCTTAAACCTGCACTCCACCAACAGCACAGCAGCAGCAGCTCCACG	60							
Db	15011	CAAAGCCGACCTTAAACCTGCACTCCACCAACAGCACAGCAGCAGCAGCTCCACG	15070							
Qy	61	TCTGACGGAGGCGCTCTTCGGCTCCCGCGCGGCCACTCGCTCCCGCTGGCGAGACGGT	120							
Db	15071	TCTGACGGAGGCGCTCTTCGGCTCCCGCGCGGCCACTCGCTCCCGCTGGCGAGACGGT	15130							
Qy	121	CGTGTTGAGCCCTATGTGGACATTGTGAGTTTACCGGCTCTGGAGCGTGGACCATGGC	180							
Db	15131	CGTGTTGAGCCCTATGTGGACATTGTGAGTTTACCGGCTCTGGAGCGTGGACCATGGC	15190							
Qy	181	GACACAGCGTGTGACAGACCG	204							
Db	15191	GACACAGCGTGTGACAGACCG	15214							

```

RESULT 10
US-09-925-300-330
; Sequence 330, Application US/09925300
; Patent NO. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 330
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (643)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (657)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (685)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-330

```

```

RESULT 11
US-09-918-995-32946
US-09-918-995-32946
; Sequence 32946, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32946
; LENGTH: 409
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-32946

Query Match      97.6%; Score 199.2; DB 11; Length 409;
Best Local Similarity 98.5%; Pred. No. 4.4e-50;
Matches 201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1  CAAAGCCCGACCTTAAACCTGTGAGTCACCAACGACGACGACGACGAGCAGCTCCAGC 60
Db      48  CAAAGCCCGACCTTAAACCTGTGAGTCACCAACGACGACGACGAGCAGCTCCAGC 107

Qy      61  TCTGACGGAGGCGCTTTCGCGTCCCGGCCGCCACATCGCTCCCGCTGGCGAGGACGT 120
Db      108  TCTGACGGAGGCGCTTTCGCGTCCCGGCCGCCACATCGCTCCCGCTGGCGAGGACGT 167

Qy      121  CGTGTTGAGCCCTATGTGGAATTGTGAGTTTACCGCTCTGTGAGCGTGGACCATGGC 180
Db      168  CGTGTTGAGCCCTATGTGGAATTGTGAGCTTTACCGCTCTGTGAGCGTGGACCATGGC 227

Qy      181  GAGCAGAGCGTGTGACGACCG 204
Db      228  GAGCAGAGCGGTTGACGACCG 251

```

```

RESULT 12
US-09-866-582-17
; Sequence 17, Application US/09866582
; Patent No. US20020127620A1
; GENERAL INFORMATION:
; APPLICANT: Witman, George B.
; APPLICANT: Pazour, Gregory J.
; APPLICANT: Rosenbaum, Joel L.
; APPLICANT: Cole, Douglas G.
; TITLE OF INVENTION: INTRAFACELLULAR TRANSPORT
; FILE REFERENCE: 07917-145001
; CURRENT APPLICATION NUMBER: US/09/866,582
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 60/206,923
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 2036
; TYPE: DNA
; ORGANISM: Chlamydomonas reinhardtii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1517)
US-09-866-582-17

```

```
Db 812 CTGCATCAACTGGCGCGCCCTTCATCCGCTCTCTCGTCACTTCGAGCACCTGCCCGT 871
Qy 111 CGAGGAGCGTGTGTGAGCCCTATGTGACTTTGTGCTAGTTTACGGCTCTGGAGCGT 170
Db 872 GGTGGAGTTTGTAGCTGGAGCGGGCGTGTGACGACGAGGCGGGCGGCTGCTGGGCGA 931
Qy 171 GGACCATGGCGAGCAG 186
Db 932 GGACGCGGCATGGAG 947

RESULT 13
US-10-101-464A-289
; Sequence 289, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR FILING DATE: 1999-01-12
; PRIOR FILING DATE: 2000-11-01
; PRIOR FILING DATE: 1999-01-12
; PRIOR FILING DATE: 1999-11-01
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 289
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-289

Query Match 17.8%; Score 36.4; DB 14; Length 591;
Best Local Similarity 50.6%; Pred. No. 0.19;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 30 CAACACGACACGACGAGCAGCTCAGCTCTGACGAGGCGCTCTTCGCTCCCGGCC 89
Db 364 CAAAGTCCACCCCGGAGCTTCGACGCTCCGTCGAGGCGACCCCTCGTCTTTCAGCTGGCG 423
Qy 90 CGGCCACTCGCTCCCGCTGGCGAGGACGCTGCGTGTGAGCCCTATGTGACTTTGCTGA 149
Db 424 CTCGCCCTGCGCTCCCGACCTCGCGGCGGCGGCGCTTCTTCCGACCTTTCGC 483
Qy 150 GTTTTACCGCTCTGGAGCGTGGACCATGCGGAGCAGAGCGTGTGACAGCACC 203
Db 484 CTTTCGTCGCGACGGGAGCGCGAGCTGCTCTTCTTACAGCATGCGGCGGCTGG 537

RESULT 14
US-10-101-464A-864
; Sequence 864, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR FILING DATE: 1999-01-12
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 289
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-864

Query Match 17.8%; Score 36.4; DB 14; Length 591;
Best Local Similarity 50.6%; Pred. No. 0.19;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 30 CAACACGACACGACGAGCAGCTCAGCTCTGACGAGGCGCTCTTCGCTCCCGGCC 89
Db 364 CAAAGTCCACCCCGGAGCTTCGACGCTCCGTCGAGGCGACCCCTCGTCTTTCAGCTGGCG 423
Qy 90 CGGCCACTCGCTCCCGCTGGCGAGGACGCTGCGTGTGAGCCCTATGTGACTTTGCTGA 149
Db 424 CTCGCCCTGCGCTCCCGACCTCGCGGCGGCGGCGGCGCTTCTTCCGACCTTTCGC 483
Qy 150 GTTTTACCGCTCTGGAGCGTGGACCATGCGGAGCAGAGCGTGTGACAGCACC 203
Db 484 CTTTCGTCGCGACGGGAGCGGACGCTGCTCTTCTTACAGCATGCGGCGGCTGG 537

RESULT 15
US-10-156-761-1749
; Sequence 1749, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1749
; LENGTH: 2922
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2922)
US-10-156-761-1749

Query Match 17.7%; Score 36.2; DB 14; Length 2922;
Best Local Similarity 49.7%; Pred. No. 0.24;
Matches 92; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy 11 CTTTACCTTGGAGTCCACCAACGACGACGAGCAGCAGTCCAGCTCTGACGAG 70
Db 194 CTTGGACCTGTGAGGCGGACCGACCTGCGCGAGTACACCAAGTCAGCATGCCCTGGG 253
Qy 71 GCCTTTCCGCTCCCGCGCCGCCACCTCGCTCCGCGTGGCGAGGAGCGTGTGAGC 130
Db 254 GAGAGTTCCCGCGCGCTGCGCGCGGACCGGCGGCTGTACGAGCGCGAGTGG 313
```

```
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 864
; LENGTH: 2380
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-864

Query Match 17.8%; Score 36.4; DB 14; Length 2380;
Best Local Similarity 50.6%; Pred. No. 0.21;
Matches 88; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy 30 CAACACGACACGACGAGCAGCTTCAGCTCTGACGAGGCGCTTCGCTCCCGGCC 89
Db 364 CAAAGTCCACCCCGGAGCTTCGACGCTCCGTCGAGGCGACCCCTCGTCTTTCAGCTGGCG 423
Qy 90 CGGCCACTCGCTCCCGCTGGCGAGCAGCTGCTGTGAGCCCTATGTGACTTTGCTGA 149
Db 424 CTCGCCCTGCGCTCCCGACCTCGCGGCGGCGGCGCTTCTTCCGACCTTTCGC 483
Qy 150 GTTTTACCGCTCTGGAGCGTGGACCATGCGGAGCAGAGCGTGTGACAGCACC 203
Db 484 CTTTCGTCGCGACGGGAGCGCGAGCTGCTCTTCTTACAGCATGCGGCGGCTGG 537
```

Qy	131	CCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGGACCA	CGGCGGAGAGCG	190
Db	314	ACGTCCCGCCGAGTGGCCGACCGCGGTGCTCTCCGGTGGGCC	CGGCGGAGAGCG	373
Qy	191	TGGTG		195
Db	374	TGCTG		378

Search completed: December 4, 2003, 09:00:10  
Job time : 314.892 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:08:45 ; Search time 23.2027 Seconds  
(without alignments)  
465.179 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSPITLTLQSTNTHQSSSSS.....AEFYELWSVDHGEQSWVTP 68

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-embl/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2001.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2002.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseq-embl/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	359	100.0	504	18	AAW26706 Human TAB1 (TAK1 b
2	359	100.0	504	18	AAW26707 Human TAB1 (TAK1 b
3	359	100.0	504	20	AAV09541 Human TAB1 protein
4	359	100.0	504	21	AAV91001 Human TAB-1 protein
5	359	100.0	504	21	AAV93450 Human TAB1 protein
6	359	100.0	513	20	AAV09550 Human TAB1 protein
7	359	100.0	516	24	ABU11578 Human MDDT polypep
8	359	100.0	517	20	AAV09546 Human TAB1-FLAG pr
9	352	98.1	84	21	AAB56692 Human prostate can

10	77	21.4	16	20	AAV09549 Human TAB1 peptide
11	70	19.5	1162	22	ABB70838 Drosophila melanog
12	69.5	19.4	199	22	AAW77804 Human HT4SG64 seri
13	68.5	19.1	113	22	AAU47075 Propionibacterium
14	67.5	18.8	739	22	ABG16477 Novel human diago
15	67.5	18.8	739	22	ABG16477 Novel human diago
16	67.5	18.8	739	24	ABR41097 Atrophin-1 interac
17	67.5	18.8	739	24	AAE32720 Human atrophin-1 i
18	67.5	18.8	852	20	AAV30948 Atrophin-1 interac
19	66	18.4	16	20	AAV09548 Human TAB1 peptide
20	64.5	18.0	116	22	AAU32515 Novel human secret
21	64.5	18.0	244	21	AAV86525 Human gene 72-enco
22	63.5	17.7	187	21	AAV86525 Arabidopsis thalia
23	63.5	17.7	219	21	AAV86525 Arabidopsis thalia
24	63.5	17.7	240	19	AAW5243 Fragment of human
25	63.5	17.7	240	23	AAE27007 Human gene 13 enco
26	63.5	17.7	240	23	AAE27007 Human gene 13 enco
27	63.5	17.7	240	24	ABU65018 Human secreted pro
28	63.5	17.7	241	22	ABG03626 Novel human diago
29	63.5	17.7	243	21	AAV35653 Arabidopsis thalia
30	63.5	17.7	335	19	AAW5208 Human secreted pro
31	63.5	17.7	335	23	AAE26971 Human gene 13 enco
32	63.5	17.7	335	23	AAE26971 Human gene 13 enco
33	63.5	17.7	335	24	ABU64982 Human secreted pro
34	63.5	17.7	391	22	AAV36176 Human D1E2 protein
35	63.5	17.7	824	22	AAV20251 Human ADAM8. Homo
36	63.5	17.7	824	23	AAU86156 Human PRO1686 poly
37	63.5	17.7	10431	24	ABU54861 Human CA125 amino
38	63	17.5	66	22	AAV86430 Human immune/haema
39	63	17.5	2112	22	ABB60403 Drosophila melanog
40	62.5	17.4	92	21	AAV2825 Human ORF2589
41	62.5	17.4	675	22	AAU30204 Propionibacterium
42	62.5	17.4	1945	22	ABB64947 Drosophila melanog
43	62	17.3	527	19	AAW48419 Amino acid sequenc
44	62	17.3	727	24	ABB84668 Human SECP-20 prot
45	62	17.3	734	22	AAV82316 Human immunoglobul

## ALIGNMENTS

### RESULT 1

AAW26706

ID AAW26706 standard; Protein; 504 AA.

XX AC AAW26706;

XX AC AAW26706;

DT 25-MAR-2003 (updated)

DT 14-APR-1998 (first entry)

XX DT 14-APR-1998 (first entry)

DE Human TAB1 (TAK1 binding protein).

XX DE Human TAB1 (TAK1 binding protein).

XX TAB1; TAK1 binding protein; transforming growth factor-beta;

XX signal transduction; human.

XX OS Homo sapiens.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Misc-difference 52

XX FT Misc-difference 52 /note= "variant has Arg as residue 52"

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

XX FT Misc-difference 52

102a



XX (CHUS ) CHUGAI SEIYAKU KK.  
 PA Ohtomo T, Ono K, Tsuchiya M;  
 PI WPI; 1999-312645/26.  
 XX N-PSDB; AAY56278.  
 PT Screening for TGF- beta inhibitory substances, which are useful as  
 PT drugs for treatment of diseases relating to its disorder  
 PS Claim 3; Page 147-149; 195pp; Japanese.  
 XX A method has been developed for screening for substances which inhibit  
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method  
 CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 CC and (b) detecting the amount of bound polypeptide, in which the sample  
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming  
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
 CC indications e.g. as TGF-beta signal transduction inhibitors or  
 CC activators, or extracellular matrix protein production enhancement  
 CC inhibitors or activators, or cell proliferation prevention inhibitors or  
 CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence represents human TAB1.  
 XX Sequence 504 AA;  
 SQ

Query Match 100.0%; Score 359; DB 20; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-37;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 60  
 |||||  
 Db 437 QSPFTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 496

QY 61 EQSVVTAP 68  
 |||||  
 Db 497 EQSVVTAP 504

RESULT 4  
 AAY91001  
 ID AAY91001 standard; Protein; 504 AA.  
 XX  
 AC AAY91001;  
 XX  
 DT 04-SEP-2000 (first entry)  
 XX  
 DE Human TAB-1 protein sequence SEQ ID NO:4.  
 XX  
 KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;  
 KW screening; signal transduction; inhibition; inflammatory cytokine;  
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;  
 KW antiinflammatory; suppression.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200023610-A1.  
 XX  
 XX 27-APR-2000.  
 XX  
 XX 21-OCT-1999; 99WO-JP05817.  
 XX  
 XX 21-OCT-1998; 98JP-0299962.  
 XX  
 XX (CHUS ) CHUGAI SEIYAKU KK.  
 XX  
 XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;  
 XX

DR WPI; 2000-339707/29.  
 DR N-PSDB; AAA39106.  
 XX  
 PT Method for screening inhibitors of TAK1 signal transduction for  
 PT suppression of inflammatory cytokine production and use as  
 PT antiinflammatory agents  
 XX  
 PS Disclosure; Page 90-94; 100pp; Japanese.  
 XX  
 CC The present invention describes a method for screening compounds for  
 CC inhibition of inflammatory cytokine signal transduction by contacting  
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition  
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds  
 CC for inhibition of inflammatory cytokine signal transduction in which the  
 CC inhibition of inflammatory cytokine signal transduction is used as a  
 CC compositions for the treatment of inflammatory disorders containing as  
 CC active component an inflammatory cytokine signal transduction inhibitor.  
 CC TAK1 is an essential component of the signalling process which results  
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),  
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used  
 CC for the selection of effective antiinflammatory agents. The present  
 CC sequence represents human TAB-1, which is used in the exemplification of  
 CC the present invention.  
 XX Sequence 504 AA;  
 SQ

Query Match 100.0%; Score 359; DB 21; Length 504;  
 Best Local Similarity 100.0%; Pred. No. 3.6e-37;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 60  
 |||||  
 Db 437 QSPFTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLMSVDHG 496

QY 61 EQSVVTAP 68  
 |||||  
 Db 497 EQSVVTAP 504

RESULT 5  
 AAY59450  
 ID AAY59450 standard; Protein; 504 AA.  
 XX  
 AC AAY59450;  
 XX  
 DT 24-MAR-2000 (first entry)  
 XX  
 DE Human TAB1 protein sequence.  
 XX  
 KW Human; TAB1; XIAP; X-linked inhibitor of apoptosis protein; TGF-beta;  
 KW transforming growth factor-beta activated kinase 1; monocyte migration;  
 KW TAK1 binding protein 1; extracellular matrix protein production;  
 KW cell growth inhibitor; beta-amyloid protein deposition;  
 KW immunosuppression; Transforming growth factor-beta.  
 XX  
 OS Homo sapiens.  
 XX  
 XX JP11326328-A.  
 XX  
 XX 26-NOV-1999.  
 XX  
 XX 13-MAY-1998; 98JP-0130378.  
 XX  
 XX 13-MAY-1998; 98JP-0130378.  
 XX  
 XX (MATS/) MATSUMOTO K.  
 XX  
 XX WPI; 2000-078337/07.  
 DR N-PSDB; AAZ48861.  
 XX  
 XX Screening a substance which inhibits combination of the X-linked  
 PT inhibitor of apoptosis protein -  
 XX

PS	Claim 2; Page 25-26; 43pp; Japanese.	
XX	This sequence represents the human TAB1 protein.	
CC	The invention relates to a method for screening a substance inhibiting	
CC	the formation of a complex between XIAP and TAB1, in which X-linked	
CC	inhibitor of apoptosis protein (XIAP), transforming growth factor-beta	
CC	activated kinase 1 (TAK1) binding protein 1 (TAB1) and a substance to be	
CC	tested are contacted with each other and then the presence or formation	
CC	of a complex between XIAP and TAB1 is detected. The substance can be used	
CC	as a drug for extracellular matrix protein production enhancement, cell	
CC	growth inhibition, monocyte migration, physiologically active substance	
CC	induction, immunosuppression, and beta-amyloid protein deposition. A	
CC	substance inhibiting the formation of a complex between TAB1 and XIAP as	
CC	well as between XIAP and TGF-beta (Transforming growth factor-beta) type	
CC	I and/or type II receptor is useful as a drug.	
XX		
Sequence	504 AA;	
Query Match	100.0%; Score 359; DB 21; Length 504;	
Best Local Similarity	100.0%; Pred. No. 3.6e-37;	
Matches	68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QSPFTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVPEYVDFAEFYRLWSVDHG 60	
DB	437 QSPFTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVPEYVDFAEFYRLWSVDHG 496	
QY	61 EQSVVTAP 68	
DB	497 EQSVVTAP 504	
RESULT 6		
AAAY09550		
ID	AAAY09550 standard; Protein; 513 AA.	
AC	AAAY09550;	
XX		
DT	21-JUL-1999 (first entry)	
XX		
DE	Human TAB1 protein SEQ ID NO:43.	
XX		
KW	Human; TAB1; TAK1; screening; inhibition; TGF-beta;	
KW	transforming growth factor beta.	
OS	Homo sapiens.	
XX		
PN	WO9921010-A1.	
XX		
PD	29-APR-1999.	
XX		
PF	22-OCT-1998; 98WO-JP04796.	
XX		
PR	22-OCT-1997; 97JP-0290188.	
XX		
PA	(CHUS ) CHUGAI SEIYAKU KK.	
XX		
PI	Ohtomo T, Ono K, Tsuchiya M;	
XX		
DR	WPI; 1999-312645/26.	
DR	N-PSDB; AAX56310.	
XX		
PT	Screening for TGF- beta inhibitory substances, which are useful as	
PT	drugs for treatment of diseases relating to its disorder	
XX		
PS	Example 13; Page 186-188; 195pp; Japanese.	
XX		
CC	A method has been developed for screening for substances which inhibit	
CC	the binding of TAK1 polypeptide to TAB1 polypeptide. The method	
CC	comprises: (a) contacting the polypeptide in the presence of a sample;	
CC	and (b) detecting the amount of bound polypeptide, in which the sample	
CC	can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming	
CC	growth factor (TGF)-beta inhibitory substances can be used in drugs for	
CC	indications e.g. as TGF-beta signal transmission inhibitors or	

CC	activators, or extracellular matrix protein production enhancement	
CC	inhibitors or activators; or cell proliferation prevention inhibitors or	
CC	activators, or monocyte migration inhibitors or activators, or	
CC	physiological activity induction inhibitors or activators, or	
CC	immunosuppression inhibitors or activators, or amyloid beta protein	
CC	precipitation inhibitors or activators, and such substances can also be	
CC	inhibitors of the TAK1 polypeptide function, particularly kinase	
CC	activity. The present sequence represents human TAB1.	
XX		
Sequence	513 AA;	
Query Match	100.0%; Score 359; DB 20; Length 513;	
Best Local Similarity	100.0%; Pred. No. 3.6e-37;	
Matches	68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 QSPFTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVPEYVDFAEFYRLWSVDHG 60	
DB	446 QSPFTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEGGRVPEYVDFAEFYRLWSVDHG 505	
QY	61 EQSVVTAP 68	
DB	506 EQSVVTAP 513	
RESULT 7		
ABU11578		
ID	ABU11578 standard; Protein; 516 AA.	
XX		
AC	ABU11578;	
XX		
DT	12-FEB-2003 (first entry)	
XX		
DE	Human MDDT polypeptide SEQ ID 525.	
XX		
KW	MDDT; human; disease detection and treatment molecule polypeptide;	
KW	anti-inflammatory; immunosuppressive; osteopathic; cytostatic; anti-HIV;	
KW	haemostatic; nephrotropic; antianaemic; antipsoriatic; hepatotropic;	
KW	gene therapy; protein replacement therapy; cell proliferative disorder;	
KW	cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; sarcoma;	
KW	anaemia; Crohn's disease; acquired immunodeficiency syndrome; AIDS;	
KW	Geopasture's syndrome; inflammation; osteoporosis; thrombocytopaenia;	
XX	psoriasis; hepatitis.	
OS	Homo sapiens.	
XX		
PN	WO200279449-A2.	
XX		
PD	10-OCT-2002.	
XX		
PF	27-MAR-2002; 2002WO-US09944.	
XX		
PR	28-MAR-2001; 2001US-279619P.	
PR	29-MAR-2001; 2001US-280067P.	
PR	29-MAR-2001; 2001US-280068P.	
PR	16-MAY-2001; 2001US-291280P.	
PR	17-MAY-2001; 2001US-291829P.	
PR	17-MAY-2001; 2001US-291849P.	
PR	19-JUN-2001; 2001US-293428P.	
PR	20-JUN-2001; 2001US-299776P.	
PR	20-JUN-2001; 2001US-300001P.	
XX		
PA	(INCY-) INCYTE GENOMICS INC.	
XX		
PI	Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;	
PI	Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ameshey SR;	
PI	Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleeefeld Y, Gerstin EH;	
PI	Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;	
PI	Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;	
XX		
DR	WPI; 2003-058431/05.	
DR	N-PSDB; ABX34568.	
XX		
PT	New purified disease detection and treatment molecule proteins and	



PT polynucleotides, useful for diagnosing, treating or preventing cancers  
 PT (e.g. leukemia or sarcoma), anemia, Crohn's disease, AIDS, osteoporosis  
 PT or hepatitis

PS Claim 27; SEQ ID NO 525; 339pp + Sequence Listing; English.

XX This invention describes a novel disease detection and treatment molecule  
 CC polypeptide (MDDT) which has anti-inflammatory, immunosuppressive,  
 CC osteopathic, cytostatic, anti-HIV, haemostatic, nephrotropic,  
 CC antianemic, antipsoriatic and hepatotropic activity. The polynucleotides  
 CC and the polypeptides of the invention can be used for gene therapy,  
 CC protein replacement therapy and are useful for treating a variety of  
 CC diseases or conditions. These polypeptides or polynucleotides are  
 CC particularly useful for diagnosing, treating or preventing cell  
 CC proliferative disorders (e.g. cancers including adenocarcinoma,  
 CC leukaemia, lymphoma, melanoma, myeloma or sarcoma), anaemia, Crohn's  
 CC disease, acquired immunodeficiency syndrome (AIDS), Goodpasture's  
 CC syndromes, inflammation, osteoporosis, thrombocytopaenia, psoriasis or  
 CC hepatitis. ABU11450-ABU11845 represent the MDDT polynucleotides encoded  
 CC by ABU11450-ABU11845, described in the disclosure of the invention.  
 CC NOTE: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 516 AA;

Query Match 100.0%; Score 359; DB 24; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-37;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPILTLQSTNTHQTQSSSSDGGGLFRSRPAHSLPPGEGRVEPYVDFAEFRLMSVDHG 60  
 |||||  
 DB 449 QSPILTLQSTNTHQTQSSSSDGGGLFRSRPAHSLPPGEGRVEPYVDFAEFRLMSVDHG 508  
 |||||  
 QY 61 EQSVVTAP 68  
 |||||  
 DB 509 EQSVVTAP 516  
 |||||

RESULT 8

AA09546  
 ID AAY09546 standard; Protein; 517 AA.

XX AC AAY09546;

XX 21-JUL-1999 (first entry)

XX Human TAB1-FLAG protein.

XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
 KW transforming growth factor beta.

XX Homo sapiens.

OS Synthetic.

XX WO9921010-A1.

XX 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS ) CHUGAI SEIYAKU KK.

XX Ohtomo T, Ono K, Tsuchiya M;

XX WPI; 1999-312645/26.

XX N-PSDB; AAX56282.

XX Screening for TGF- beta inhibitory substances, which are useful as  
 PT drugs for treatment of diseases relating to its disorder

XX

PS Example 1; Page 163-166; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit  
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method  
 CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 CC and (b) detecting the amount of bound polypeptide, in which the sample  
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming  
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
 CC indications e.g. as TGF-beta signal transduction inhibitors or  
 CC activators, or extracellular matrix protein production enhancement  
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or  
 CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence represents TAB1-FLAG from an example of  
 CC the present invention.

XX Sequence 517 AA;

Query Match 100.0%; Score 359; DB 20; Length 517;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-37;  
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPILTLQSTNTHQTQSSSSDGGGLFRSRPAHSLPPGEGRVEPYVDFAEFRLMSVDHG 60  
 |||||  
 DB 437 QSPILTLQSTNTHQTQSSSSDGGGLFRSRPAHSLPPGEGRVEPYVDFAEFRLMSVDHG 496  
 |||||  
 QY 61 EQSVVTAP 68  
 |||||  
 DB 497 EQSVVTAP 504  
 |||||

RESULT 9

AA056692  
 ID AAB56692 standard; Protein; 84 AA.

XX AC AAB56692;

XX 13-MAR-2001 (first entry)

XX Human prostate cancer antigen protein sequence SEQ ID NO:1270.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;  
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;  
 KW vulnery; gastrointestinal; nephrotropic; antiinfective; gynaecological;  
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;  
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;  
 KW wound; infectious disease.

XX Homo sapiens.

XX WO200055174-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05988.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM;

XX WPI; 2000-587513/55.

XX N-PSDB; AAF15895.

XX Prostate cancer associated gene sequences, referred to as prostate  
 PT cancer antigens, useful for treatment, prevention, and diagnosis of  
 PT disorders such as prostate cancer

XX

PS Claim 11; Page 1693; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytostatic,

CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,

CC nephrotropic, antineoplastic, gynaecological and antibacterial activities,

CC and can be used in gene therapy. The prostate cancer antigen, chromosome

CC polynucleotides may be used for detection of prostate cancer, diagnostic

CC identification, as chromosome markers, and for numerous other diagnostic

CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

CC AAB57303 represent sequences used in the exemplification of the present

XX invention.

SQ Sequence 84 AA;

Query Match 98.1%; Score 352; DB 21; Length 84;

Best Local Similarity 98.5%; Pred. No. 2.7e-37;

Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QSPFTLQSTNTHTQSSSSDGLFRSPAHSLPGDGRVPEYVDFAEFYRLWSVDHG 60

DB 17 QSPFTLQSTNTHTQSSSSDGLFRSPAHSLPGDGRVPEYVDFAEFYRLWSVDHG 76

QY 61 EQSVVTAP 68

DB 77 EQSVVTAP 84

RESULT 10

AAV09549

ID AAY09549 standard; peptide; 16 AA.

XX AC AAY09549;

XX 21-JUL-1999 (first entry)

XX Human TAB1 peptide TAB1C-2.

XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;

XX transforming growth factor beta.

XX Homo sapiens.

XX WO9921010-A1.

XX 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Ohtomo T, Ono K, Tsuchiya M;

XX WPI; 1999-312645/26.

XX Screening for TGF- beta inhibitory substances, which are useful as

XX drugs for treatment of diseases relating to its disorder

XX Example 13; Page 182; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit

XX the binding of TAK1 polypeptide to TAB1 polypeptide. The method

XX comprises: (a) contacting the polypeptide in the presence of a sample;

XX and (b) detecting the amount of bound polypeptide, in which the sample

XX can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming

XX growth factor (TGF)-beta inhibitory substances can be used in drugs for

XX indications e.g. as TGF-beta signal transmission inhibitors or

CC activators, or extracellular matrix protein production enhancement

CC inhibitors or activators, or cell proliferation prevention inhibitors or

CC activators, or monocyte migration inhibitors or activators, or

CC physiological activity induction inhibitors or activators, or

CC immunosuppression inhibitors or activators, or amyloid beta protein

CC precipitation inhibitors or activators, and such substances can also be

CC inhibitors of the TAK1 polypeptide function, particularly kinase

CC activity. The present sequence represents a peptide from an example

XX of the present invention.

SQ Sequence 16 AA;

Query Match 21.4%; Score 77; DB 20; Length 16;

Best Local Similarity 100.0%; Pred. No. 0.0054;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSPFTLQSTNTHTQ 15

DB 2 QSPFTLQSTNTHTQ 16

RESULT 11

ABB70838

ID ABB70838 standard; Protein; 1162 AA.

XX AC ABB70838;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 39306.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL14941.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions -

XX Disclosure; SEQ ID NO 39306; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins

XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1162 AA;

SQ Query Match 19.5%; Score 70; DB 22; Length 1162;

CC quantitate the presence of similar nucleic acids in samples. The PSPase  
CC polypeptides may be used as antigens in the production of antibodies  
CC against the PSPase polypeptides and in assays to identify modulators of  
CC PSPase expression and activity. The anti-PSPase antibodies and  
CC antagonists may also be used to down regulate expression and activity,  
CC the anti-pspase antibodies may also be used as diagnostic agents for  
CC detecting the presence of pspase polypeptides in samples. Disorders that  
CC may be prevented, diagnosed and/or treated by the invention are:  
CC Immune/autoimmune disorders (e.g. Wiscott-Aldrich syndrome,  
CC Chediak-Higashi syndrome, Hashimoto's thyroiditis and multiple  
CC sclerosis); inflammatory conditions (e.g. Crohn's disease, inflammatory  
CC bowel disease, appendicitis and rheumatoid arthritis); cellular  
CC proliferative disorders (e.g. lymphoma, lung and intestinal cancers); and  
CC cardiovascular disorders (e.g. Scimitar syndrome, Ebstein's anomaly and  
CC aneurysm).

XX

SQ Sequence 199 AA;

Query Match 19.4%; Score 69.5; DB 22; Length 199;  
Best Local Similarity 41.9%; Pred. No. 1.3;  
Matches 18; Conservative 3; Mismatches 17; Indels 5; Gaps 1

QY 1 QSPILTQSTNTHQTSSSSSDGG----LFRSRPAHSLPAGE 38  
||||| : ||||| :  
Db 119 QSPILTQDRIHRSRDAPSSISGCSKFTAVRKMADKLFPVG 161  
||||| : ||||| :

RESULT 13

AAU47075

ID AAU47075 standard; Protein; 113 AA.

XX AC AAU47075;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #7971.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.  
PR 02-JUN-2000; 2000US-208841P.  
PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX SKelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
PI N-PSDB; AAS59536.

XX WPI: 2001-616774/71.  
XX N-PSDB; AAS59536.

XX PT Propionibacterium acnes polypeptides and nucleic acids useful for  
PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -

XX Example 1; SEQ ID NO 8270; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.

CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies  
 CC specific for P. acnes proteins. These antibodies can be used to  
 CC downregulate expression and activity of P. acnes polypeptides and  
 CC therefore treat P. acnes infections. The antibodies may also be used as  
 CC diagnostic agents for determining P. acnes presence, for example, by  
 CC enzyme linked immunosorbent assay (ELISA).  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 113 AA;  
 SQ  
 Query Match 19.1%; Score 68.5; DB 22; Length 113;  
 Best Local Similarity 35.4%; Pred. No. 0.86;  
 Matches 17; Conservative 8; Mismatches 22; Indels 1; Gaps 1;  
 QY 3 PTLTQSTNTHQTQSSSSDGLFRSPAHSLPDPEDGRVPEYVDFAE 50  
 DB 61 PTLTAATSGPQWNTPSGT-GGLVGPSPMKSTPPHDHGIWTPRFRAD 107

RESULT 14  
 ABGL6477  
 ID ABGL6477 standard; Protein; 739 AA.  
 XX  
 AC ABGL6477;  
 DT 18-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #16468.  
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 KW Homo sapiens.  
 OS  
 XX WO200175067-A2.  
 FN  
 XX 11-OCT-2001.  
 PD  
 XX 30-MAR-2001; 2001WO-US08631.  
 PF  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Dmanac RT, Liu C, Tang YT;  
 PI  
 XX WPI; 2001-639362/73.  
 DR N-PSDB; AAS80664.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 XX Claim 20; SEQ ID No 46836; 103pp; English.  
 PS  
 XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 739 AA;  
 SQ  
 Query Match 18.8%; Score 67.5; DB 22; Length 739;  
 Best Local Similarity 31.1%; Pred. No. 13;  
 Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;  
 QY 2 SPTLTQSTNTHQTQSSSSDGLFRSPAHSLPDPEDGRVPEYVDF 48  
 DB 119 SSTGSLPPTNTNTSEGATSGLLIPLTISGGSPRLNPTOALPPGWEQVDF 173  
 QY 49 AEFVRLMSVDHGQ 62  
 DB 174 -QHGRVYVDHVEK 186

RESULT 15  
 AAG68173  
 ID AAG68173 standard; Protein; 739 AA.  
 XX  
 AC AAG68173;  
 DT 25-JAN-2002 (first entry)  
 XX  
 DE Atrophin-1 interacting protein (AIP4) SEQ ID NO:89.  
 XX Human; high bone mass; HEM gene; Zmax1 gene; chromosome 11; 11q13.3;  
 KW sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;  
 KW antisense therapy; vaccine; bone disorder; Paget's disease;  
 KW sclerostosis; osteomalacia; fibrous dysplasia.  
 XX Homo sapiens.  
 OS  
 XX WO200177327-A1.  
 FN  
 XX 18-OCT-2001.  
 PD  
 XX 21-JUN-2000; 2000WO-US16951.  
 PF  
 XX 05-APR-2000; 2000US-0543771.  
 PR 05-APR-2000; 2000US-0544398.  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PA  
 XX Carulli JP, Little RD, Recker RR, Johnson ML;  
 PI  
 XX WPI; 2001-657171/75.  
 XX  
 XX New high bone mass (HEM) and Zmax1 genes and proteins useful for  
 PT modulating bone mass for the treatment of e.g. osteoporosis -  
 PT  
 XX Claim 76; Page 392-394; 443pp; English.  
 PS  
 XX The present invention describes the human Zmax1 gene and the high bone  
 CC mass (HEM) gene which are found on chromosome 11q13.3. The Zmax1 and  
 CC HEM genes have osteopathic activities. The genes can be used in gene  
 CC therapy, antisense therapy and in the production of vaccines. They  
 CC can be used in the diagnosis and treatment of bone disorders including  
 CC osteoporosis, Paget's disease, sclerostosis, osteomalacia and fibrous  
 CC dysplasia. ABA82038 to ABA82700 and AAG68168 to AAG68193 represent



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 07:04:49 ; Search time 57.3851 Seconds  
(without alignments)  
3783.871 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091

Perfect score: 1252

Sequence: 1 gtagagcttgcgagttatc.....cattacagatcctgtgcag 684

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09830144/runat\_03122003\_122343\_21194/app\_query.fasta\_1.1230  
-DB=A\_Geneseq\_19Jun03 -QFWT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09830144 @CGN 1 1 0 @runat\_03122003\_122343\_21194 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFBUCK=100 -JONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_19Jun03.\*

1: /SIDS1/gcgdata/geneq/geneq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/geneq/geneq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/geneq/geneq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/geneq/geneq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/geneq/geneq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/geneq/geneq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/geneq/geneq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/geneq/geneq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/geneq/geneq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/geneq/geneq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/geneq/geneq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/geneq/geneq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/geneq/geneq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/geneq/geneq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/geneq/geneq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/geneq/geneq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/geneq/geneq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/geneq/geneq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/geneq/geneq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/geneq/geneq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/geneq/geneq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/geneq/geneq-emb1/AA2001.DAT.\*  
23: /SIDS1/gcgdata/geneq/geneq-emb1/AA2002.DAT.\*  
24: /SIDS1/gcgdata/geneq/geneq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1252	100.0	567	20	AA28998	Human TGF-beta act
2	1252	100.0	579	18	AAW27092	Mouse transforming
3	1252	100.0	579	18	AAW27093	Human transforming
4	1252	100.0	579	20	AA28996	Human TGF-beta act
5	1252	100.0	579	20	AA28996	Human TAK1 protein
6	1252	100.0	579	21	AA28996	Human TAK1 protein
7	1252	100.0	579	23	AA28996	Human TAK1-1 protei
8	1252	100.0	579	23	AA28996	Pain regulated pro
9	1252	100.0	579	20	AA28996	Human TAK1-6xHis p
10	700	55.9	678	22	AB58061	Human TGF-beta act
11	414	33.1	252	22	AB58061	Drosophila melanog
12	374	29.9	367	21	AB58061	Drosophila melanog
13	374	29.9	369	21	AB58061	Arabidopsis thalia
14	374	29.9	407	21	AB58061	Arabidopsis thalia
15	373	29.8	369	21	AB58061	Arabidopsis thalia
16	373	29.8	374	21	AB58061	Arabidopsis thalia
17	373	29.8	412	21	AB58061	Arabidopsis thalia
18	371	29.6	349	22	AB58061	Human colon cancer
19	371	29.6	455	21	AB58061	A human regulator
20	371	29.6	455	21	AB58061	Human survival reg
21	371	29.6	455	21	AB58061	A human cardiovasc
22	371	29.6	473	22	AB58061	Human protein sequ
23	371	29.6	800	22	AB58061	Human TGF-beta rec
24	371	29.6	800	22	AB58061	Novel protein kina
25	371	29.6	800	23	AB58061	Motif zipper conta
26	363.5	29.0	394	24	AB58061	Human protein kina
27	363.5	29.0	1020	22	AB58061	Drosophila melanog
28	359.5	28.7	1024	23	AB58061	Human NOV7, a mixe
29	355.5	28.4	1021	23	AB58061	Novel human protei
30	355	28.4	719	22	AB58061	Human protein kina
31	355	28.4	1036	23	AB58061	Novel human protei
32	355	28.4	1036	24	AB58061	Amino acid sequenc
33	355	28.4	1036	24	AB58061	Human kinase and p
34	353.5	28.2	859	16	AB58061	Human leucine zipp
35	353.5	28.2	859	18	AB58061	Human leucine-zipp
36	353.5	28.2	859	24	AB58061	Polypeptide sequen
37	352.5	28.2	888	23	AB58061	Mouse ischaemic co
38	351.5	28.1	977	22	AB58061	Drosophila melanog
39	350	28.0	589	21	AB58061	Arabidopsis thalia
40	350	28.0	732	21	AB58061	Arabidopsis thalia
41	350	28.0	760	21	AB58061	Arabidopsis thalia
42	349.5	27.9	1046	22	AB58061	Human kinase (PKIN
43	349.5	27.9	1097	23	AB58061	Human PKIN-12 prot
44	348	27.8	319	21	AB58061	Arabidopsis thalia
45	348	27.8	341	21	AB58061	Arabidopsis thalia

# ALIGNMENTS

## RESULT 1

AA28998  
ID AA28998 standard; Protein; 567 AA.

XX AA28998;

XX AC (first entry)

DT 29-OCT-1999

XX Human TGF-beta activated kinase (TAK) 1c amino acid sequence.

XX Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;

XX TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;

XX intractable disease; atrophic dermatitis; psoriasis; viral infection;

XX endotoxin shock; septicemia; human; hTAK1c.

XX Homo sapiens.

XX OS

XX WO9940202-A1.

PN

XX PD 12-AUG-1999.  
XX PF 02-FEB-1999; 99WO-JP00422.  
XX PR 30-OCT-1998; 98JP-0309316.  
XX PR 06-FEB-1998; 98JP-0026003.  
XX PA (TANA) TANABE SEIYAKU CO.  
XX PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;  
XX DR WPI; 1999-494298/41.  
XX DR N-PSDB; AAX99698.  
XX PT Nuclear factor kappa B activation inhibitors, useful as preventives  
XX PT for, e.g. autoimmune diseases  
XX PS Examples; Page 43-46; 49pp; Japanese.  
XX CC The invention provides a method for identifying or screening a nuclear  
XX CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a  
XX CC test substance on modulating the function(s) of TGF-beta activated kinase  
XX CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
XX CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
XX CC intractable diseases with inflammation (such as atrophic dermatitis and  
XX CC psoriasis), viral infection, endotoxin shock, septicemia and others. The  
XX CC present sequence represents the amino acid sequence of human TAK1c  
XX CC (hTAK1c) protein.  
XX SQ Sequence 567 AA;  
  
Alignment Scores:  
Pred. No.: 7,32e-134 Length: 567  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0  
  
US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAY28998 (1-567)  
  
QY 1 GTGAGCTTCGGAGTATCCGCTGCAACATCCCTTAATTTGTAAGCTTTATGAGCC 60  
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyAla 95  
  
QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTCAAGGGGCTCTTATATATG 120  
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115  
  
QY 121 CTGCATGGTCTGAACCATTCCTATATTAATGCTGCTGCCCAAGCAATGAGTTGGTTTA 180  
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 135  
  
QY 181 CAGTGTCTCCAGAGTGGCTTACTCTCAGATGCAACCCAAAGCGCTAAATTCACAGG 240  
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg 155  
  
QY 241 GACCTGAACACCAAACTTACTCTGCTGCTGAGGGGGAGCAGTTCTAAATAATTTGTAT 300  
Db 156 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysLeuCysAsp 175  
  
QY 301 TTGTTGTCAGCTGTGACATTCAGACACACATGACCAATACAGGGAGTGGCTTGG 360  
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195  
  
QY 361 ATGCACCTCAAGTTTGTAGGTAGTAATACAGTGAATAATGTCAGCTTTCAGCTGG 420  
Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 215  
  
QY 421 GGTATTATCTTTGGGAGTGATACCGCTCGGAACCTTTGATGATGAGTGGGCCA 480  
Db 216 GlyIleIleLeuTyrGluValIleThrArgLysProPheAspGluIleGlyGlyPro 235

QY 481 GCTTTCCGATCATGCTGGCTGTTTCATATGCTACTCGACCACTGATAAAAAATTTA 540  
Db 236 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255  
  
QY 541 CCTAAGCCCATTTGAGAGCCCTGATGACTCGTTTGTCTTAAAGATCCTTCCAGCCCT 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 275  
  
QY 601 TCAATGGAGAAATTTGTAATATGACTCACTTGTATCGGTACTTTCAGGAGCAGAT 660  
Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
  
QY 661 GAGCCATTACAGTATCTTGTCTGAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303  
  
RESULT 2  
AAW27092  
ID AAW27092 standard; Protein; 579 AA.  
XX AC AAW27092;  
XX DT 19-NOV-1997 (first entry)  
XX DE Mouse transforming growth factor-beta activated kinase TAK-1.  
XX KW TGF-beta; signal transmission; TGF-beta activated kinase;  
XX KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;  
XX KW protein kinase.  
XX OS Mus musculus.  
XX JN JP09163990-A.  
XX PD 24-JUN-1997.  
XX PF 27-SEP-1996; 96JP-0256747.  
XX PR 24-JUL-1996; 96US-0685625.  
XX PR 29-SEP-1995; 95JP-0253549.  
XX PA (CHUS) CHUGAI PHARM CO LTD.  
XX PA (UENO) UENO N.  
XX DR WPI; 1997-380171/35.  
XX DR N-PSDB; AAT85094.  
XX PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1  
XX PT - useful for studying the TGF-beta signal transmission system  
XX PS Claim 14; Page 10-12; 20pp; Japanese.  
XX CC The present sequence represents mouse transforming growth factor-beta  
XX CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the  
XX CC TAK-1 protein which is involved in the TGF-beta family signal  
XX CC transmission system. TAK-1, also known as activator of MAPK Kinase  
XX CC (AMK-1), is an enzyme which is activated by TGF-beta and bone  
XX CC morphogenetic protein (BMP) and activates MAPK kinase by  
XX CC phosphorylation.  
XX SQ Sequence 579 AA;  
  
Alignment Scores:  
Pred. No.: 7,38e-134 Length: 579  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 18 Gaps: 0  
  
US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAW27092 (1-579)



QY 1 GTAGAGCTCGCAGTATCCCGTGTGAACCATCTTAATATTGAAAGCTTTATGAGCC 60  
 Db ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95  
 QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATAAATGTG 120  
 Db CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115  
 QY 121 CTGCATCGTGTGAACCATCTCCCATATTATATCTGCTGCCACGCAATGAGTGTGTTA 180  
 Db LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 135  
 QY 181 CAGTGTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 240  
 Db GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155  
 QY 241 GACCTGAAACCAACTTACTCTGCTGCTGAGGGGAGCAGTCTAAATAATTTGTGAT 300  
 Db AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
 QY 301 TTTGGTACAGCTGTGACATTCACACACATGACCAATACACAGGGGAGTGTCTGG 360  
 Db PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195  
 QY 361 ATGCACCTCAAGTTTTTGAAGTAGTAATTACAGTGAATAATGTGACGCTTCAGCTGG 420  
 Db MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 215  
 QY 421 GGTATATTCTTTGGGAAGTGATTAACGGTCTGGAACCTTTGATGAGATTGTGGCCCA 480  
 Db GlyIleLeuLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
 QY 481 GCTTTCCGAATCATGTGGCTGTTCAATATGTTACTGACCATCGACCATGATAAAATTTA 540  
 Db AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255  
 QY 541 CCTAAGCCCATTCAGAGCCTGATGCTGTGTTGCTTAAAGATCCTTCCAGCGCCCT 600  
 Db ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 275  
 QY 601 TCAATGGAGAAATTTGAAATAATGACTCATCTGATGGCGTACTTTCCAGGAGCAGAT 660  
 Db SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
 QY 661 GAGCCATTACAGTATCTCTCTCAG 684  
 Db GluProLeuGlnTyrProCysGln 303

## RESULT 3

AAW27093  
 ID AAW27093 standard; Protein; 579 AA.

XX AC AAW27093;

XX DT 19-NOV-1997 (first entry)

XX DE Human transforming growth factor-beta activated kinase TAK-1.

XX DE TGF-beta; signal transduction; TGF-beta activated kinase;

XX KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;

XX KW protein kinase.

XX OS Homo sapiens.

XX PN JF09163990-A.

XX PD 24-JUN-1997.

XX PF 27-SEP-1996; 96JP-0256747.

XX PR 24-JUL-1996; 96US-0685625.

XX PR 29-SEP-1995; 95JP-0253549.

XX

PA (CHUS ) CHUGAI PHARM CO LTD.  
 PA (UENO/) UENO N.  
 XX  
 DR WPI; 1997-380171/35.  
 DR N-ESDB; AAT85095.  
 XX  
 PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1  
 PT - useful for studying the TGF-beta signal transduction system  
 XX  
 FS Claim 15; Page 13-15; 20pp; Japanese.

XX The present sequence represents human transforming growth factor-beta  
 CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the  
 CC TAK-1 protein which is involved in the TGF-beta family signal  
 CC transduction system. TAK-1, also known as activator of MAPK Kinase  
 CC (AMK-1), is an enzyme which is activated by TGF-beta and bone  
 CC morphogenetic protein (BMP) and activates MAPK kinase by  
 CC phosphorylation.

XX Sequence 579 AA;

## Alignment Scores:

Pred. No.: 7.38e-134 Length: 579  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 18 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAW27093 (1-579)

QY 1 GTAGAGCTCGCAGTATCCCGTGTGAACCATCTTAATATTGAAAGCTTTATGAGCC 60  
 Db ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95  
 QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATAAATGTG 120  
 Db CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115  
 QY 121 CTGCATCGTGTGAACCATCTCCCATATTATATCTGCTGCCACGCAATGAGTGTGTTA 180  
 Db LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 135  
 QY 181 CAGTGTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 240  
 Db GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155  
 QY 241 GACCTGAAACCAACTTACTCTGCTGCTGAGGGGAGCAGTCTAAATAATTTGTGAT 300  
 Db AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
 QY 301 TTTGGTACAGCTGTGACATTCACACACATGACCAATACACAGGGGAGTGTCTGG 360  
 Db PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195  
 QY 361 ATGCACCTCAAGTTTTTGAAGTAGTAATTACAGTGAATAATGTGACGCTTTCAGCTGG 420  
 Db MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 215  
 QY 421 GGTATATTCTTTGGGAAGTGATTAACGGTCTGGAACCTTTGATGAGATTGTGGCCCA 480  
 Db GlyIleLeuLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
 QY 481 GCTTTCCGAATCATGTGGCTGTTCAATATGTTACTGACCATCGACCATGATAAAATTTA 540  
 Db AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255  
 QY 541 CCTAAGCCCATTCAGAGCCTGATGCTGTGTTGCTTAAAGATCCTTCCAGCGCCCT 600  
 Db ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 275  
 QY 601 TCAATGGAGAAATTTGAAATAATGACTCATCTGATGGCGTACTTTCCAGGAGCAGAT 660

Db 276 SerMetGluGluIleValIleMetThrHisLeuMetArgTyr:PheProGlyAlaAsp 295

QY 661 GAGCCATTACAGTATCCTTGTGTCAG 684

Db 296 GluProLeuGlnTyr:ProCysGln 303

# RESULT 4

ID AAY28996 standard; Protein; 579 AA.

AC AAY28996;

XX 29-OCT-1999 (first entry)

DT Human TGF-beta activated kinase (TAK) 1a amino acid sequence.

XX Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;  
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;  
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;  
KW endotoxin shock; septicemia; human; hTAK1a.

XX Homo sapiens.

OS WO9940202-A1.

PN 12-AUG-1999.

PD 02-FEB-1999; 99WO-JP00422.

PF 30-OCT-1998; 98JP-0309316.

PR 06-FEB-1998; 98JP-0026003.

PA (TANA) TANABE SEIYAKU CO.

XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;

PI WPI; 1999-494298/41.

DR N-PSDB; AAX39696.

XX Nuclear factor kappa B activation inhibitors, useful as preventives  
PT for, e.g. autoimmune diseases

XX Examples; Page 35-39; 49pp; Japanese.

XX The invention provides a method for identifying or screening a nuclear  
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a  
CC test substance on modulating the function(s) of TGF-beta activated kinase  
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
CC intractable diseases with inflammation (such as atrophic dermatitis and  
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The  
CC present sequence represents the amino acid sequence of human TAK1a  
CC (hTAK1a) protein.

XX Sequence 579 AA;

## Alignment Scores:

Pred. No.: 7,38e-134 Length: 579  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAY28996 (1-579)

QY 1 GTACAGCTCGCGAGTATCCGCTGTAACCATCTTAATATGTAAAGCTTTATGAGCC 60

Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIleLeuTyr:GlyAla 95

QY 61 TGCTTGATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120

Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115

QY 121 CTGCGATGGTGGTGAACCATTTGCCATATTACTGCTGCCGCCCGCAATGAGTTGGTGTTA 180  
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 135  
QY 181 CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 240  
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155  
QY 241 GACCTGAAACCAACCAACTTACTGCTGGTTCAGGGGGGACAGTCTTAAATAATTCGTGAT 300  
Db 156 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
QY 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATAACAAGGGGAGTGTGCTTGG 360  
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195  
QY 361 ATGGCACCTGAAAGTGTGAAAGTAGTAATTACAGTGAAGAAATGTGACGTCTTCAGCTGG 420  
Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
QY 421 GGTATTATTCTTTGGGAAGTGATACGCGTCCGAAACCCCTTGATGAGATTGTTGGCCCA 480  
Db 216 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
QY 481 GCTTTCCGAATCATGTGGGCTGTTCATAATGTGTACTCGACACCACCTGTATATAAATTTA 540  
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255  
QY 541 CCTAAGCCCATTCAGAGCCCTGATGACTCGTTGTGTCTTAAGATCCTTCCCGAGCGCCT 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGGAGAAATTCGAAATAATGACTCACTGATCGGTACTTCCAGGAGCAGAT 660  
Db 276 SerMetGluGluIleValIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
QY 661 GAGCCATTACAGTATCCTTGTGTCAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303

## RESULT 5

AAY09542

ID AAY09542 standard; Protein; 579 AA.

XX AC AAY09542;

XX 21-JUL-1999 (first entry)

DT Human TAK1 protein.

XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;  
XX transforming growth factor beta.

XX Homo sapiens.

PH Key Location/Qualifiers  
FT CDS 183...1922 /\*tag= a

XX WO9921010-A1.

XX 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Ohtomo T, Ono K, Tsuchiya M;

XX WPI; 1999-312645/26.

DR N-PSDB; AAX56279.

PT Screening for TGF- $\beta$  inhibitory substances, which are useful as

PT drugs for treatment of diseases relating to its disorder

XX

PS Claim 4; Page 155-157; 195pp; Japanese.

XX

CC A method has been developed for screening for substances which inhibit

CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method

CC comprises: (a) contacting the polypeptide in the presence of a sample;

CC and (b) detecting the amount of bound polypeptide, in which the sample

CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming

CC growth factor (TGF)- $\beta$  inhibitory substances can be used in drugs for

CC indications e.g. as TGF- $\beta$  signal transduction inhibitors or

CC activators, or extracellular matrix protein production enhancement

CC inhibitors or activators, or cell proliferation prevention enhancement

CC activators, or monocyte migration inhibitors or activators, or

CC physiological activity induction inhibitors or activators, or

CC immunosuppression inhibitors or activators, or amyloid  $\beta$  protein

CC precipitation inhibitors or activators, and such substances can also be

CC inhibitors of the TAK1 polypeptide function, particularly kinase

CC activity. The present sequence represents human TAK1.

XX

SQ Sequence 579 AA;

Alignment Scores:

Pred. No.:	7.38e-134	Length:	579
Score:	1252.00	Matches:	228
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	20	Gaps:	0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAY09542 (1-579)

QY	1	GTAGAGCTCGCGAGTATCCGTTGAGAACCATCTTAATATTGTAAAGCTTTATGGAGCC	60
DB	76	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla	95
QY	61	TGCTTGATCCAGTGTCTTGATGAATATGCTGAAGGGGCTCTTTATATAATGTG	120
DB	96	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal	115
QY	121	CTGCATGTGTGACCAATTCCTATATTATCTGCGCCAGCAATGAGTGTGTGTTA	180
DB	116	LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu	135
QY	181	CAGTGTTCCTCCAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG	240
DB	136	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg	155
QY	241	GACCTGAACCAACCAACTTACTCTGTTGTCAGGGGGGACAGTCTTAAATAATTGTGAT	300
DB	156	AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp	175
QY	301	TTTGGTCAGCTGTGACATTCAGACACATGACCAATACAGGGAGGCTGCTGG	360
DB	176	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr	195
QY	361	ATGCACCTGAAGTTTGTGAAGGTAGTAATTACAGTGAATAATGTGACGCTTCAGCTGG	420
DB	196	MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr	215
QY	421	GGTATTATCTTTGGGAGTATACCGCTCGAAACCCCTTTGATGAGATTGGTGCCCA	480
DB	216	GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro	235
QY	481	GCTTTCGAATCATGTGGCTGTTTCAATAAGTACTGACCCACCACTCATATAAATTTA	540
DB	236	AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu	255
QY	541	CCTAAGCCCAATTGAGAGCTGATGACTGCTGTGTGGTCTTAAGATCTCTCCAGCGCCT	600

DB	256	ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro	275
QY	601	TCAATGAGGAAATGTGAAATATGACTCACTGATGCGGTACTTCCAGGACAGAT	660
DB	276	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp	295
QY	661	GAGCCATTACAGTATCCCTTGTCTCAG	684
DB	296	GluProLeuGlnTyrProCysGln	303

RESULT 6

AAAY91000

ID AAY91000 standard; Protein; 579 AA.

XX AAY91000;

AC AAY91000;

XX

DT 04-SEP-2000 (first entry)

XX

DE Human TAK-1 protein sequence SEQ ID NO:2.

XX

KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;

KW screening; signal transduction; inhibition; inflammatory cytokine;

KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;

KW antiinflammatory; suppression.

XX

OS Homo sapiens.

XX

PN WO200023610-A1.

XX

PD 27-APR-2000.

XX

XX 21-OCT-1999; 99WO-JP05817.

XX

PR 21-OCT-1998; 98JP-0299962.

XX

PA (CHUS) CHUGAI SEIYAKU KK.

XX

PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;

XX

DR WPI; 2000-339707/29.

DR N-PSDB; AAA39105.

XX

PT Method for screening inhibitors of TAK1 signal transduction for

PT suppression of inflammatory cytokine production and use as

PT antiinflammatory agents -

XX

PS Example 1; Page 80-84; 100pp; Japanese.

XX

CC The present invention describes a method for screening compounds for

CC inhibition of inflammatory cytokine signal transduction by contacting

CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition

CC of TAK1/TAB1 binding. Also described is a method for screening compounds

CC for inhibition of inflammatory cytokine signal transduction in which the

CC inhibition of TAK1 phosphorylation is selected for; and drug

CC compositions for the treatment of inflammatory disorders containing as

CC active component an inflammatory cytokine signal transduction inhibitor.

CC TAK1 is an essential component of the signalling process which results

CC in release of inflammatory cytokines such as interleukin-1 (IL-1),

CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used

CC for the selection of effective antiinflammatory agents. The present

CC sequence represents human TAK-1, which is used in the exemplification of

CC the present invention.

XX

SQ Sequence 579 AA;

Alignment Scores:

Pred. No.:	7.38e-134	Length:	579
Score:	1252.00	Matches:	228
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAY91000 (1-579)

QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGCAACCACTTAATATTCTTAAGCTTTATGGAGCC 60  
 Db 76 ValGluLeuArgGlnLeuSerArgValAlaHisProAsnIleValIysLeuTyrGlyAla 95  
 QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
 Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115  
 QY 121 CTGCATGGTCTGAACCACTTACTATATATCTGCTGCCACGCAATGAGTGTGTGTTA 180  
 Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
 QY 181 CAGTGTTCACAGGAGTGGCTTCTTCTCAGCATGCAACCAAGGCTTAATTCACAGG 240  
 Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 155  
 QY 241 GACCTGAAACCAACCACTTACTGCTGTGCTGCGGGGGGACAGTCTTAAAAATTTGTGAT 300  
 Db 156 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
 QY 301 TTGGTACACCTGTGATTCAGACACACATGACCAATACAGAGGGAGTGTGCTGG 360  
 Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195  
 QY 361 ATGCACCTGACGTTTTGTGAGGTAGTAATACAGTGAATAAATGTGACGCTTCAGCTGG 420  
 Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
 QY 421 GGTATTATCTTTGGGAGTGATTAACCGTGGGAACCTTTGATGAGATGTGTGGCCCA 480  
 Db 216 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
 QY 481 GCTTTCGGAATCATGTGGGCTGTTCATAATGTGACTGACACACACCTGATAAAAAATTA 540  
 Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255  
 QY 541 CCTAAGCCCATTTGAGAGCTGATGACGCTGTGTGTGTTGTTAAAGATCCTCCAGCGCCT 600  
 Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
 QY 601 TCATGAGGAGAAATGTGAAATAATGACTCACCTTGTATGGGTTACTTTCCAGGAGCAGAT 660  
 Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
 QY 661 GAGCCATTACAGTATCCTTGTGAC 684  
 Db 296 GluProLeuGlnTyrProCysGln 303

## RESULT 7

ABB85033

ID ABB85033 standard; Protein; 579 AA.

AC ABB85033;

DT 16-MAY-2002 (first entry)

DE Pain regulated protein sequence 28.

XX Pain; analgesic; gene therapy; neurological disorder;  
 KW neurodegenerative disease.  
 XX Homo sapiens.  
 OS

XX WO200212338-A2.

XX 14-FEB-2002.

XX 03-AUG-2001; 2001WO-EP09011.

XX 03-AUG-2000; 2000DE-1037759.

XX

(CHEF ) GRUENENTHAL GMBH.

XX Gallen C, Wetzels I, Wnendt S, Weihe E, Schaefer MK;

XX WPI; 2002-257469/30.

DR N-PSDB; ABL88437.

XX Identifying pain-regulating compounds, useful for treating chronic pain  
 PT and for diagnosis, by measuring binding of compounds to specific  
 PT peptides and proteins

XX Claim 1; Fig 44; 213pp; German.

XX The invention relates to identifying pain-regulating substances (A)  
 CC comprises (i) incubating a test substance with a cell (or preparation  
 CC from it) that has synthesised a peptide or protein (B) and (ii) measuring  
 CC either binding of the test substance to (B) or some functional parameter  
 CC that is altered by this binding. The method is useful for identifying  
 CC pain-regulating substances (A) with analgesic activity. (A) along with  
 CC nucleic acid (ABL88411-ABL88441) that encode proteins (B,  
 CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the  
 CC nucleic acid; antibodies against (B); cells that express (B) and agents  
 CC that bind to (B), are all useful for treating pain, particularly chronic  
 CC pain, including use in gene therapy. The same materials can also be used  
 CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The  
 CC present sequence is that of a polypeptide of the invention.

XX Sequence 579 AA;

SQ

Alignment Scores:  
 Pred. No.: 7,38e-134 Length: 579  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservations: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 23 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x ABB85033 (1-579)

QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGCAACCACTTAATATTCTTAAGCTTTATGGAGCC 60  
 Db 76 ValGluLeuArgGlnLeuSerArgValAlaHisProAsnIleValIysLeuTyrGlyAla 95  
 QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
 Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115  
 QY 121 CTGCATGGTCTGAACCACTTACTATATATCTGCTGCCACGCAATGAGTGTGTGTTA 180  
 Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
 QY 181 CAGTGTTCACAGGAGTGGCTTCTTCTCAGCATGCAACCAAGGCTTAATTCACAGG 240  
 Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 155  
 QY 241 GACCTGAAACCAACCACTTACTGCTGTGCTGCGGGGGGACAGTCTTAAAAATTTGTGAT 300  
 Db 156 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
 QY 301 TTGGTACACCTGTGATTCAGACACACATGACCAATACAGAGGGAGTGTGCTGG 360  
 Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195  
 QY 361 ATGCACCTGACGTTTTGTGAGGTAGTAATACAGTGAATAAATGTGACGCTTCAGCTGG 420  
 Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
 QY 421 GGTATTATCTTTGGGAGTGATTAACCGTGGGAACCTTTGATGAGATGTGTGGCCCA 480  
 Db 216 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
 QY 481 GCTTTCGGAATCATGTGGGCTGTTCATAATGTGACTGACACACACCTGATAAAAAATTA 540

Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255  
 QY 541 CCTAAGCCCATGAGAGCTGATGACTGTTGTTGCTTAAGATCCTTCCAGGCGCCT 600  
 Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
 QY 601 TCAATGAGGAAATGTGAAATATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 660  
 Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
 QY 661 GAGCATTACAGTATCCTTGTGTCAG 684  
 Db 296 GluProLeuGlnTyrProCysGln 303

RESULT 8  
 AAY09547  
 ID AAY09547 standard; Protein; 590 AA.  
 XX  
 AC AAY09547;  
 XX  
 XX 21-JUL-1999 (first entry)  
 DT  
 DE Human TAK1-6xHis protein.  
 XX  
 XX Human; TAK1; TAK1; screening; inhibition; TGF-beta;  
 KW transforming growth factor beta.  
 KW  
 XX Homo sapiens.  
 OS Synthetic.  
 OS  
 XX WO9921010-A1.  
 PN  
 XX 29-APR-1999.  
 PD  
 XX 22-OCT-1998; 98WO-JP04796.  
 PF  
 XX 22-OCT-1997; 97JP-0290188.  
 PR  
 XX (CHUS ) CHUGAI SEIYAKU KK.  
 PA  
 XX Ohtomo T, Ono K, Tsuchiya M;  
 XX  
 XX WPI; 1999-312645/26.  
 DR N-PSDB; AAX56285.  
 DR  
 XX Screening for TGF- beta inhibitory substances, which are useful as  
 FT drugs for treatment of diseases relating to its disorder  
 PT  
 XX Example 1; Page 171-174; 195pp; Japanese.

CC A method has been developed for screening for substances which inhibit  
 CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method  
 CC comprises: (a) contacting the polypeptide in the presence of a sample;  
 CC and (b) detecting the amount of bound polypeptide, in which the sample  
 CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming  
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for  
 CC indications e.g. as TGF-beta signal transmission inhibitors or  
 CC activators, or extracellular matrix protein production enhancement  
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or  
 CC activators, or monocyte migration inhibitors or activators, or  
 CC physiological activity induction inhibitors or activators, or  
 CC immunosuppression inhibitors or activators, or amyloid beta protein  
 CC precipitation inhibitors or activators, and such substances can also be  
 CC inhibitors of the TAK1 polypeptide function, particularly kinase  
 CC activity. The present sequence represents TAK1-6xHis from an example of  
 CC the present invention.

SQ Sequence 590 AA;  
 Alignment Scores:  
 Pred. No.: 7,43e-134 Length: 590  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0  
 US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAY09547 (1-590)  
 QY 1 GTAGAGCTTCGGCAGTATCCGTTGTGAACCATCTATATTTAAAGCTTTATCGACC 60  
 Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95  
 QY 61 TGCTTGAATCCAGTGTGTCTTGTGATGGAATATCTGAAGGGGGCTCTTATATAATGTG 120  
 Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115  
 QY 121 CTGATGTGTGTAACCATTCCTATATATCTGCTGCCACCGCAATGAGTTGGTGTTA 180  
 Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
 QY 181 CAGTGTTCCTCCAGGAGTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAGG 240  
 Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155  
 QY 241 GACCTGAAACCCACCAACTTACTGCTGTTGAGGGGGACAGTTCTAAAAATTTGTGAT 300  
 Db 156 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
 QY 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATACAAAGGGAGTGCTGCTTGG 360  
 Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195  
 QY 361 ATGGCAGCTGAAGTTTGTGAGGTAGTAAATTTACAGTGAATAATGTGAGCTTTCAGCTGG 420  
 Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
 QY 421 GGTATTATTCTTTGGGAAGTGATAACCGTCGGAACCCCTTTGATGAGATTGGTGGCCCA 480  
 Db 216 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
 QY 481 GCTTTCCGAATCATGTGGGTGTTTCATAAATGGTACTCGACCCACCTGATATAAAATTTA 540  
 Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255  
 QY 541 CCTAAGCCCATGAGAGCTGATGACTGTTGTTGTTAAGATCCTTCCAGGCGCCT 600  
 Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
 QY 601 TCAATGAGGAAATGTGAAATATGACTCACTTGTGCGGTACTTTCCAGGAGCAGAT 660  
 Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
 QY 661 GAGCATTACAGTATCCTTGTGTCAG 684  
 Db 296 GluProLeuGlnTyrProCysGln 303

RESULT 9  
 AAY28997  
 ID AAY28997 standard; Protein; 606 AA.  
 XX  
 AC AAY28997;  
 XX  
 XX 29-OCT-1999 (first entry)  
 DT  
 DE Human TGF-beta activated kinase (TAK) 1b amino acid sequence.  
 XX Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;  
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;  
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;  
 KW endotoxin shock; septicemia; human; hTAK1b.  
 XX  
 OS Homo sapiens.  
 OS  
 XX WO9940202-A1.  
 PN  
 XX

PD 12-AUG-1999.  
 XX  
 PF 02-FEB-1999; 99WO-JP00422.  
 XX  
 PR 30-OCT-1998; 98JP-0309316.  
 PR 06-FEB-1998; 98JP-0026003.  
 XX  
 XX (TANA) TANABE SEIYAKU CO.  
 PA  
 Hasegawa K, Kageyama N, Sakurai H, Sugita T;  
 XX  
 WPI; 1999-494298/41.  
 DR N-PSDB; AAX39697.  
 XX  
 XX Nuclear factor kappa B activation inhibitors, useful as preventives  
 PT for, e.g. autoimmune diseases  
 XX  
 PS Examples; Page 39-43; 49pp; Japanese.  
 XX  
 CC The invention provides a method for identifying or screening a nuclear  
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a  
 CC test substance on modulating the function(s) of TGF-beta activated kinase  
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
 CC intractable diseases with inflammation (such as atrophic dermatitis and  
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The  
 CC present sequence represents the amino acid sequence of human TAK1B  
 CC (hTAK1B) protein.  
 XX  
 XX Sequence 606 AA;

Alignment Scores:  
 Pred. No.: 7,51e-134 Length: 606  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0  
 US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAY28997 (1-606)

QY 1 GTAGAGCTTCGGCAGTTCATCCGCTGCAACCATCTAATATTTAAAGCTTTATGAGCC 60  
 Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyVala 95  
 QY 61 TGCTTGAATCCAGTGTCTGTGTGATGGAATATCTGCAAGGGGCTCTTTATATATATG 120  
 Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115  
 QY 121 CTGCATGCTGCTGAACCATTCCTATATATCTGCTGCCACGCAAGAGTGTGTGTTA 180  
 Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
 QY 181 CAGTGTTCACCAAGGAGTGGCTATCTTTCACAGATGCAACCAAGCGCTTAATTCACAGG 240  
 Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArg 155  
 QY 241 GACCTGAAACCAACCAACTTACTGCTGCTGGTTCAGGGGGCAGTCTTAAATAATTTGAT 300  
 Db 156 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
 QY 301 TTTGGTACAGCTGCTGACATTCACACACATGACCATCAACAGGGGAGTGTGCTGG 360  
 Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195  
 QY 361 ATGCACCTGAAGTGTGAGGTAGTAAATACAGTGAATAATGACGCTCTTCAGCTGG 420  
 Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215  
 QY 421 GGTATATCTTTGGGAGTGAATACCGCTCGGAACCTTTGATGAGATGTTGGGCCCA 480  
 Db 216 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235

QY 481 GCTTTCGAATCATGTGGGCTGTTTCAATGTTACTCGACCCACTGATAAAAAATTTA 540  
 Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeu 255  
 QY 541 CCTAAGCCCATTCAGAGCCCTGATGACTCGTGTGTGGTCTAAAGATCTTCCAGCGCCCT 600  
 Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
 QY 601 TCAATGAGGAAATTTGAAAAATATGACTCATCTGATCGGTACTTCCAGGAGCAGAT 660  
 Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
 QY 661 GAGCCATTACAGTATCCTGTCTAG 684  
 Db 296 GluProLeuGlnTyrProCysGln 303  
 RESULT 10  
 ABB58061  
 ID ABB58061 standard; Protein; 678 AA.  
 XX  
 AC ABB58061;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SEQ ID NO 975.  
 XX  
 XX Drosophila; developmental biology; cell signalling; insecticide;  
 XX pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 XX 23-MAR-2000; 2000US-191637P.  
 PR 11-JUL-2000; 2000US-0614150.  
 XX  
 XX (PEKE) PE CORP NY.  
 PA  
 Venter JC, Adams M, Li PWD, Myers EW;  
 PI  
 WPI; 2001-656860/75.  
 DR N-PSDB; ABL02164.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SEQ ID NO 975; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (ABB57737-ABB72072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published/pct\_sequences.  
 XX  
 SQ Sequence 678 AA;  
 Alignment Scores:  
 Pred. No.: 9.5e-71 Length: 678  
 Score: 700.00 Matches: 130  
 Percent Similarity: 73.01% Conservative: 35  
 Best Local Similarity: 57.52% Mismatches: 57  
 Query Match: 55.91% Indels: 4

```

DB: 22 Gaps: 3
US-09-830-144-1_COPY_408_1091 (1-684) x ABB58061 (1-678)
QY 4 GAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTTGAAGCTTTATGGA----- 57
Db 60 GluValLysGlnLeuSerArgValLysHisProAsnIleAlaLeuHisGlyIleSer 79
QY 58 GCCTGCTGAATCCAGTGTCTTGTGATGAATATGTCGAAGGGGCTCTTTATATAAT 117
Db 80 SerTyrGlnGlnAlaThrTyrLeuIleMetGluPheAlaGluGlyGlySerLeuHisAsn 99
QY 118 GTGCTCATGCTGCTGACCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 177
Db 100 PheLeuHisGly---LysValLysProAlaTyrSerLeuAlaHisAlaMetSerTrpAla 118
QY 178 TTACAGTGTTCCTCAAGAGTGGCTTATCTTCAGCATGCAACCCCAAGCGCTTAATTCAC 237
Db 119 ArgGlnCysAlaGluGlyLeuAlaTyrLeuHisAlaMetThrProLysProLeuIleHis 138
QY 238 AGGACCTGAACACCAACAACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 297
Db 139 ArgAspValLysProLeuAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 158
QY 298 GATTTTGGTACACCTGTGCATTCACACACATGACCAATCAACAGGGGAGTGTGCT 357
Db 159 AspPheGlyThrValAlaAspLysSerThrMetMetThrAsnAsnArgGlySerAlaAla 178
QY 358 TGGATGCGCCTCAAGTGTGAGGAGTAGTAATTCAGTGAAGAAATGTGACGCTTCAGC 417
Db 179 TrpMetAlaProGluValPheGluGlySerLysTyrThrGluLysCysAspIlePheSer 198
QY 418 TGGGGTATTATCTTTGGGAAGTAAACCGCTGCGGAACCCCTTGAAGATGATGTGGC 477
Db 199 TrpAlaIleValLeuTrpGluValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 218
QY 478 CGAGCTTCGCAATCATGCTGGCTGTCTAATGCTACTGACACCAACACATGATAAATAAT 537
Db 219 ---AlaTyrThrIleGlnTrpLysIleTyrLysGlyGluArgProProLeuLeuThr 237
QY 538 TTACCTAAGCCCATTTGAGAGCCTGATGACCTGCTGTTGGTCTAAAGATCCTCCAGCGC 597
Db 238 CysProLysArgIleGluAspLeuMetAlaCysTrpLysThrValProGluAspArg 257
QY 598 CCTTCAATGGAGAAATGTGAAATAATGACTCATCTGATGCGGTACTTTCCAGGAGCA 657
Db 258 ProSerMetGlnTyrIleValGlyValMetHisGluIleValLysAspTyrThrGlyAla 277
QY 658 GATGAGCCATTACAGTAT 675
Db 278 AspLysAlaLeuGluTyr 283

RESULT 11
ABB60985
ID ABB60985 standard; Protein; 252 AA.
XX AC
XX AB60985;
XX DT
XX 26-MAR-2002 (first entry)
XX DE
XX Drosophila melanogaster polypeptide SEQ ID NO 9747.
XX KW
XX Drosophila; developmental biology; cell signalling; insecticide;
XX OS
XX Drosophila melanogaster.
XX PN
XX WO200171042-A2.
XX PD
XX 27-SEP-2001.
XX PF
XX 23-MAR-2001; 2001WO-US09231.
XX

PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX PA
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX N-PSDB; ABL05088.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure; SEQ ID NO 9747; 2lpp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL01840-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABBS7737-ABBS72072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 252 AA;
Alignment Scores:
Pred. No.: 3,14e-38 Length: 252
Score: 414.00 Matches: 82
Percent Similarity: 59.69% Conservative: 35
Best Local Similarity: 41.84% Mismatches: 53
Query Match: 33.07% Indels: 26
DB: 22 Gaps: 4
US-09-830-144-1_COPY_408_1091 (1-684) x ABB60985 (1-252)
QY 4 GAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTTGAAGCTTTATGAGCCTGC 63
Db 54 GluIleTyrGlnLeuThrLysAlaSerHisValAsnIleValGluLeuTyrGlyThrSer 73
QY 64 TTGAATCCAGTGTGT-----CTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db 74 ArgHisGluGlyCysAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 93
QY 118 GTGCTGATGCTGCTGAACCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 177
Db 94 PheLeuHis---AlaLysSerLysProSerTyrSerHisAlaHisAlaPheAsnTrpAla 112
QY 178 TTACAGTGTTCCTCAAGAGTGGCTTATCTTCACAGCATGCAACCCCAAGCGCTTAATTCAC 237
Db 113 HisGlnIleAlaGlnGlyIleAlaTyrLeuHisGlyMetGlnProLysAlaValIleHis 132
QY 238 AGGACCTGAACACCAACAACTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 297
Db 133 ArgAspIleLysProLeuAsnThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 152
QY 298 GATTTTGGTACAGCTGTGACATTCAGACACATGACCAATCAACAGGGGAGTGTGCT 357
Db 153 AspPheGlyThrValValAspLeuSerGlnSerIleSerCysAsnAlaGlyThrCysArg 172
QY 358 TGGATGGCAGCTGAA----- 372
Db 173 TyrLysAlaProGluValArgGluLeuPheAspPheLysSerAsnArgIleIleAsn 192
QY 373 -----GTTTGAAGGTAGTAAATTACAGTGAATAATGTGAGCTC 411
Db 193 GlnProThrGlyPheGlnLysValLeuGlnGlyAsnLysProAspGluLysCysAspVal 212
QY 412 TTCAGCTGGGGTATTATTCTTTGGGAAGTATTAACGCTGCGGAACCCCTTTGATGAGATT 471

```

```
Db      213 TyrSerTrpAlaIleThrPheTrpGluIleLeuSerArglyslubProPheGluGlnTyr 232
      472 GGTGGCCCGACTTCGATCATGTGGCTGTTCAATAAGTACTCGA 519
      233 ---AsnThrLeuPheGluLeuTyrMetAlaIleAsnGluGlyLysarg 247

RESULT 12
AAG32053
ID AAG32053 standard; Protein; 367 AA.
XX
AC AAG32053;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38597.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 11-MAY-1999; 99US-0132487.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 21-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
```



PR 13-AUG-1999;	99US-0148565.	
PR 13-AUG-1999;	99US-0148684.	
PR 16-AUG-1999;	99US-0149368.	
PR 17-AUG-1999;	99US-0149175.	
PR 18-AUG-1999;	99US-0149426.	
PR 20-AUG-1999;	99US-0149722.	
PR 20-AUG-1999;	99US-0149723.	
PR 20-AUG-1999;	99US-0149929.	
PR 23-AUG-1999;	99US-0149902.	
PR 23-AUG-1999;	99US-0149930.	
PR 25-AUG-1999;	99US-0150566.	
PR 26-AUG-1999;	99US-0150884.	
PR 27-AUG-1999;	99US-0151065.	
PR 27-AUG-1999;	99US-0151066.	
PR 27-AUG-1999;	99US-0151080.	
PR 30-AUG-1999;	99US-0151303.	
PR 31-AUG-1999;	99US-0151438.	
PR 01-SEP-1999;	99US-0151930.	
PR 07-SEP-1999;	99US-0152363.	
PR 10-SEP-1999;	99US-0153070.	
PR 13-SEP-1999;	99US-0153758.	
PR 15-SEP-1999;	99US-0154018.	
PR 16-SEP-1999;	99US-0154039.	
PR 20-SEP-1999;	99US-0154779.	
PR 22-SEP-1999;	99US-0155139.	
PR 23-SEP-1999;	99US-0155486.	
PR 24-SEP-1999;	99US-0155659.	
PR 28-SEP-1999;	99US-0156458.	
PR 29-SEP-1999;	99US-0156596.	
PR 04-OCT-1999;	99US-0157117.	
PR 05-OCT-1999;	99US-0157753.	
PR 06-OCT-1999;	99US-0157865.	
PR 07-OCT-1999;	99US-0158029.	
PR 08-OCT-1999;	99US-0158232.	
PR 12-OCT-1999;	99US-0158369.	
PR 13-OCT-1999;	99US-0159293.	
PR 13-OCT-1999;	99US-0159294.	
PR 13-OCT-1999;	99US-0159295.	
PR 14-OCT-1999;	99US-0159329.	
PR 14-OCT-1999;	99US-0159330.	
PR 14-OCT-1999;	99US-0159331.	
PR 14-OCT-1999;	99US-0159637.	
PR 14-OCT-1999;	99US-0159638.	
PR 18-OCT-1999;	99US-0159584.	
PR 21-OCT-1999;	99US-0160741.	
PR 21-OCT-1999;	99US-0160767.	
PR 21-OCT-1999;	99US-0160768.	
PR 21-OCT-1999;	99US-0160770.	
PR 21-OCT-1999;	99US-0160814.	
PR 21-OCT-1999;	99US-0160815.	
PR 22-OCT-1999;	99US-0160980.	
PR 22-OCT-1999;	99US-0160981.	
PR 22-OCT-1999;	99US-0160989.	
PR 25-OCT-1999;	99US-0161404.	
PR 25-OCT-1999;	99US-0161405.	
PR 25-OCT-1999;	99US-0161406.	
PR 26-OCT-1999;	99US-0161359.	
PR 26-OCT-1999;	99US-0161360.	
PR 26-OCT-1999;	99US-0161361.	
PR 28-OCT-1999;	99US-0161920.	
PR 28-OCT-1999;	99US-0161922.	
PR 28-OCT-1999;	99US-0161993.	
PR 29-OCT-1999;	99US-0162142.	
Alignment Scores:		
Pred. No.:	1.35e-33	367
Score:	374.00	86
Percent Similarity:	57.80%	40
Best Local Similarity:	39.45%	70
Query Match:	29.87%	22
DB:	21	9

US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAG32053 (1-367)

QY	4	GAGCTCGCAGTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGCTGC	63
Db	135	GluValSerMetLeuAlaAsnLeuLysHisProAsnIleValArgPheIleGlyAlaCys	154
QY	64	TTGAATCCAGTG-----TGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAAT	117
Db	155	ArgLysProMetValTrpCysIleValThrGluTyrAlaLysGlyGlySerValArgGln	174
QY	118	GTGCTG-----CATGGTGTCTGAACATTCCTCATATTATATCTGCTCCACGCA	165
Db	175	PheLeuThrArgArgGlnAsnArgAlaValProLeuLysLeu-----Ala	189
QY	166	ATGAGTTGGTGTTCACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAACCAAA	225
Db	190	ValLysGlnAlaLeuAspValAlaArgGlyMetAlaTyrValHisGly-----Arg	206
QY	226	GCCTAATTACAGGGACCTGAAACCAACCACTTCTGCTGGTTCAGGGGGGACAGTT	285
Db	207	AsnPheIleHisArgAspLeuLysSerAsnLeuLeuIleSerAlaAspLysSerIle	226
QY	286	CTAAATAATTGTCATTTGGTACAGCC---TGTGACATTCAGACACAC---ATGACCAAT	339
Db	227	--LysIleAlaAspPheGlyValAlaArgIleGluValGlnThrGluGlyMetThrPro	245
QY	340	AACAAGGGGAGTCTGCTTGGATGGCACCCTGAAGTTTGAAGGTAGTAAATTACAGTGAA	399
Db	246	GluThrGlyThrTyrArgTrpMetAlaProGluMetIleGlnHisArgAlaTyrAsnGln	265
QY	400	AAATGTGACGTCTTCAGCTGGGTATTATTCTTTGGAGTGTATACGCTCGGAAACCC	459
Db	266	LysValAspValTyrSerPheGlyIleValLeuTrpGluLeuIleThrGlyLeuLeuPro	285
QY	460	TTTGATGAGATTGGTGGC-----CCAGCTTTCGATCATGTGGGCTGTTCATAATGGT	513
Db	286	PheGlnAsnMetThrAlaValGlnAlaAlaPheAlaVal-----ValAsnArgGly	302
QY	514	ACTCGACCCACTGATAAAAAATTTACCTAAGCCCATTCGAGAGCTGTGATCGTGTGT	573
Db	303	ValArgProThrValProAsnAspCysLeuProValLeuSerAspIleMetThrArgCys	322
QY	574	TGGTCTAAAGATCCTTCCAGCGCCCTTCAATGGAGGAAATGTGAAAAATAATG	627
Db	323	TrpAspAlaAsnProGluValArgProCysPheValGluValValLysLeuLeu	340
RESULT 13			
AAG32052			
ID	AAG32052 standard; Protein; 369 AA.		
XX	AAG32052;		
AC	17-OCT-2000 (first entry)		
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 38596.		
DE	Protein identification; signal transduction pathway; metabolic pathway;		
XX	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX	Arabidopsis thaliana.		
OS	EPI033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
XX	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	23-MAR-1999; 99US-0123548.		
PR	25-MAR-1999; 99US-0125788.		
PR	29-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		

PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134376.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138647.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139763.  
PR 22-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0146389.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152263.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158023.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161320.  
 PR 28-OCT-1999; 99US-0161392.  
 PR 28-OCT-1999; 99US-0161393.  
 PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.: 1,36e-33 Length: 369  
 Score: 374.00 Matches: 86  
 Percent Similarity: 57.80% Conservative: 40  
 Best Local Similarity: 39.45% Mismatches: 70  
 Query Match: 29.87% Indels: 22  
 DB: 21 Gaps: 9

US-09-830-144-1\_COPY\_408\_1091 (1-684) x AAG32052 (1-369)

QY 4 GAGCTTGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGCGCTGC 63  
 DB 137 GluValSerMetLeuAlaAsnLeuLysHisProAsnIleValArgPheIleGlyAlaCys 156  
 QY 64 TTGAATCCAGT-----TGTCTTGATGCAATATGCTGAAGGGCTCTTTATATAAT 117  
 DB 157 ArgLysProMetValTrpCysIleValThrGluTyrAlaLysGlyGlySerValArgGln 176  
 QY 118 GTGCTG-----CATGTGCTGAACCATTTGCCATATTATATCTGCTGCCACGCA 165  
 DB 177 PheLeuThrArgArgGlnAsnArgAlaValProLeuLysLeu-----Ala 191  
 QY 166 ATGAGTTGGTGTTCACAGTGTCCCAAGGAGTGGCTTTATTCACAGCATGCAACCCAAA 225  
 DB 192 ValLysGlnAlaLeuAspValAlaArgGlyMetAlaTyrValHisGly-----Arg 208  
 QY 226 GCGCTAATTCACAGGACCTGAACACCAACTTACTGCTGTGAGGGGGGACAGTT 285  
 DB 209 AsnPheIleHisArgAspLeuLysSerAspAsnLeuLeuIleSerAlaAspLysSerile 228  
 QY 286 CTAATAATTTGTATTTTGTGTACAGCC---TGTGACATTCAGACACAC---ATGACCAAT 339  
 DB 229 ---LysIleAlaAspPheGlyValAlaArgIleGluValGlnThrGluGlyMetThrPro 247  
 QY 340 AACAGGGGAGTGTGCTTGGATGGCACCCTGAAGTTTGTGAAGTAGTAATACAGTGAA 399  
 DB 248 GluThrGlyThrTyrArgTrpMetAlaProGluMetIleGlnHisArgAlaTyrAsnGln 267  
 QY 400 AATGTGACGTCTCAGCTGGGTATTATTCTTTGGGAAGTGAACCGCTCGGAACCC 459  
 DB 268 LysValAspValTyrSerPheGlyIleValLeuTrpGluLeuIleThrGlyLeuLeuPro 287  
 QY 460 TTTCATGAGATTGTGGC-----CCAGCTTCCGAATCATGTGGCTGTTTCATAATGGT 513

Db 288 PheGlnAsnMetThrAlaValGlnAlaAlaPheAlaVal-----ValAsnArgGly 304  
 QY 514 ACTCGACCACTGATATAAAATTTACCTAAGCCATTGAGAGCCTGATGCTGCTGT 573  
 Db 305 ValArgProThrValProAsnAspCysLeuProValLeuSerAspIleMetThrArgCys 324  
 QY 574 TGTGCTCTAAAGATCCTTCCAGCGCCCTTCAATGAGGAGAAATTTGTAATAATG 627  
 Db 325 TrpAspAlaAsnProGluValArgProCysPheValGluValLysLeuLeu 342  
 RESULT 14  
 AAG32051  
 ID AAG32051 standard; Protein; 407 AA.  
 XX  
 AC AAG32051;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 38595.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 hybridisation assay; genetic mapping; gene expression control; promoter;  
 termination sequence.  
 KW  
 XX Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139452.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142320.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 28-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

## Alignment Scores:

Pred. No.: 1,418-33  
Score: 374.00  
Percent Similarity: 57.80%  
Best Local Similarity: 39.45%  
Query Match: 29.87%  
Length: 407  
Matches: 86  
Conservative: 40  
Mismatch: 70  
Indels: 22



```

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1.76e-33
Score: 373.00
Percent Similarity: 57.80%
Best Local Similarity: 39.91%
Query Match: 29.79%
DB: 21
US-09-830-144-1_COPY_408_1091 (1-684) x AAG22172 (1-369)

QY 4 GAGCTTCGCGAGTATCCGCTGTAACCATCCTAATATGTAAGCTTTATGGAGCTGC 63
Db 137 GluValSerMetLeuAlaPheLeuLysHisProAsnIleValArgPheIleGlyAlaCys 156
QY 64 TTGAATCCAGTG-----TGTCTTGTGATGGAATATGCTGAAGGGGGCTTTATATAAT 117
Db 157 IleLysProMetValTyrCysIleValThrGluTyrAlaLysGlyGlySerValArgGln 176
QY 118 GTGCTG-----CATGGTCTGAACCATTCCTATATTTACAGATGCACCCAAA 165
Db 177 PheLeuThrLysArgGlnAsnArgAlaValProLeuLysLeu-----Ala 191
QY 166 ATGAGTTGGTGTGTACAGTGTTCACAGGAGTGGCTTATCTTCAGATGCACCCAAA 225
Db 192 ValMetGlnAlaLeuAspValAlaArgGlyMetAlaTyrValHisGlu-----Arg 208
QY 226 GCGCTAATTCAGGGACCTGAACCAACCACTTACTGCTGTTGTCAGGGGGGACAGTT 285
Db 209 AsnPheIleHisArgAspLeuLysSerAspAsnLeuLeuIleSerAlaAspArgSerIle 228
QY 286 CTAAATAATTGTGATTTTGGTACAGCC---TGTGACATTCAGACACAC---ATGACCAAT 339
Db 229 ---LysIleAlaAspPheGlyValAlaArgIleGluValGlnThrGluGlyMetThrPro 247
QY 340 AACAGGGGAGCTGCTGCTTGCATGCACCTCAAGATTTTTTGAAGTAGTAATTACAGTCAA 399
Db 248 GluThrGlyThrTyrArgTyrMetAlaProGluMetIleGlnHisArgProTyrThrGln 267
QY 400 AATGTGACGCTTCAGCTGGGGTATTATCTTTGGGAAGTATACCGCTGGGAACCC 459
Db 268 LysValAspValTyrSerPheGlyIleValLeuTyrGluLeuIleThrGlyLeuLeuPro 287

```

```

QY 460 TTTGATGAGATTGGTGGC-----CCAGCTTTCGGAATCATGTGGCTGTTTCATAATGGT 513
Db ||| :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
288 PheGlnAsnMetThrAlaValGlnAlaAlaPheAlaVal-----ValAsnArgGly 304
QY 514 ACTCGACCCACCACTGATAAAAAATTTACCTAAGCCCATTTGAGAGCCTGATGACTCGTTGT 573
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
305 ValArgProThrValProAlaAspCysLeuProValLeuGlyGluIleMetThrArgCys 324
QY 574 TGGTCTAAAGATCCTTCCAGCGCCCTTCAANGGAGGAAATTTGTGAAAATAATG 627
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
325 TrpAspAlaAspProGluValArgProCysPheAlaGluIleValAsnLeuLeu 342

```

Search completed: December 4, 2003, 09:02:44  
Job time : 63.3851 secs

*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: December 4, 2003, 08:37:54 ; Search time 30.0405 Seconds  
(without alignments)  
4379.374 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091  
Perfect score: 1252  
Sequence: 1 gtagagcttcggcagttatc.....cattacagtcattctgtcag 684

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlh  
-O=/cnp2/1/USPTO.spool/US098301044/runat\_03122003\_122345\_21308/app\_query.fasta\_1.1230  
-DB=PIR 76 -QPMF=fastan -SUFFIX=rpr -MINWATCH=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US098301044 @CNP 1.1.35 @runat\_03122003\_122345\_21308 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR 76:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES										
Result No.	Score	Match	Length	ID	Description					
1	1252	100.0	567	2	JC5957	transforming growt				
2	1252	100.0	579	2	JC5955	transforming growt				
3	1252	100.0	606	2	JC5956	transforming growt				
4	392.5	31.3	954	1	S68178	mixed-lineage prot				
5	374	29.9	407	2	G84635	probable protein k				
6	373	29.8	412	2	T10671	protein kinase hom				
7	363.5	29.0	394	2	JU0229	mixed-lineage prot				
8	362	28.9	1338	2	T18287	protein-tyrosine k				
9	356	28.4	462	2	S29851	protein kinase 6 (				
10	355.5	28.4	370	2	T46150	protein kinase ATN				
11	353.5	28.2	668	2	JC2363	protein kinase (EC				
12	353.5	28.2	888	2	AS5318	serine/threonine p				
13	351.5	28.1	888	2	JC5399	dual leucine zippe				
14	350	28.0	546	2	D84555	probable protein k				

ALIGNMENTS

RESULT 1

JC5957  
transforming growth factor-beta activated kinase (BC 2.7.-.-) lc - human  
C;Species: Homo sapiens (man)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 17-Mar-2000  
C;Accession: JC5957  
R;Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
A;Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-inducible transcription factor.  
A;Reference number: JC5955; MUID:98153801; PMID:9480845  
A;Accession: JC5957  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-567 <SAK>  
A;Cross-references: DDBJ,AB009358  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homologous  
C;Keywords: phosphotransferase

Alignment Scores:  
Pred No.: 3.96e-73 Length: 567  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x JC5957 (1-567)

Qy	1	GTAGAGCTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC	60
Db	76	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValValLeuTyrGlyAla	95
Qy	61	TGCTTGAATCCAGTGTCTTGTCATGATATCTGACGGGCTCTTTATATAATG	120
Db	96	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal	115
Qy	121	CTGCATGTGTGACCAATTGCCCATATTATATCTGTGCTCCACGCAATGAGTGTGTGTTA	180

protein kinase ATM  
probable protein k  
hypothetical prote  
hypothetical prote  
hypothetical prote  
mixed-lineage prot  
protein kinase hom  
protein-tyrosine k  
hypothetical prote  
protein kinase ATN  
probable protein k  
hypothetical prote  
protein kinase ATN  
protein kinase hom  
probable protein k  
protein kinase ATN  
serine/threonine-p  
probable protein k  
probable protein k  
probable protein k  
protein kinases ho  
hypothetical prote  
protein-tyrosine k  
probable mitogen-a  
probable serine/th  
protein F33E2.2 [i  
hypothetical prote  
protein-tyrosine k  
protein-tyrosine k  
protein-tyrosine k  
protein-tyrosine k  
transforming prote

Db 116 LeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
QY 181 CAGTGTCTCCCAAGAGTGGCTTATCTTCACAGATGCAACCCCAAGCGCTTAATTCACAGG 240  
Db 136 GlnCysSerGlnGlyValAlaIarYrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155  
QY 241 GACCTGAAACCCCAACTTACTGCTGGTTCAGGGGGGAGCAGTCTTAAATAATTTGTGAT 300  
Db 156 AspLeuLysProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAsp 175  
QY 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGAGTGTCTGTGG 360  
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195  
QY 361 ATGGCACCTGAAGTTTTTGAAGGTAGTAATTCAGTGAAGGAGTGTCTGTGG 420  
Db 196 MetAlaProGluValPheGluGlySerAsnTrpSerGluLysCysAspValPheSerTrp 215  
QY 421 GGTATTATTCTTTGGGAAGTATACGCGTTCGGAACCCCTTTGATGAGATTGGTGGCCCA 480  
Db 216 GlyIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
QY 481 GCTTTCCGATCATGTGGCTGTTCATATGTGTACTCGACCACTGATATAAAATTTA 540  
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeu 255  
QY 541 CCTAACCCATTGAGAGCTGATGACTCGTGTGCTTAAAGATCTTCCAGAGCCCT 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGGAGAAATTGTGAAATAATGACTCACTTGTGGTCTTAAAGATCTTCCAGAGCAGAT 660  
Db 276 SerMetGluGluValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
QY 661 GAGCCATTACAGTATCTTGTCTAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303  
RESULT 2  
JC5955  
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human  
C:Species: Homo sapiens (man)  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: JC5955  
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind  
A:Reference number: JC5955; MUID:98153801; PMID:9480845  
A:Accession: JC5955  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-579 <SAK>  
A:Cross-references: DBJ:AB009356; NID:G2924623; PIDN:BAA25025.1; PID:G2924624  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C:Keywords: phosphotransferase  
Alignment Scores:  
Pred. No.: 3 94e-73 Length: 579  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0  
US-09-830-144-1\_copy\_408\_1091 (1-684) x JC5955 (1-579)  
QY 1 GTAGAGCTTCGGCAGTTATCCGTTGAAACCATCTTAATTTGTAAGCTTTATGAGCC 60  
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95  
QY 61 TGCTTGAATCAGTGTGCTTGTGATGGAATATGCTGAGGGGCTCTTTATATAATGTG 120  
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115

QY 121 CTGATGCTCTGAAACCAATTCATATTAATCTGCTGCCCAACCAATGAGTGTGTTA 180  
Db 116 LeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
QY 181 CAGTGTCTCCCAAGAGTGGCTTATCTTCACAGATGCAACCCCAAGCGCTTAATTCACAGG 240  
Db 136 GlnCysSerGlnGlyValAlaIarYrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155  
QY 241 GACCTGAAACCCCAACTTACTGCTGGTTCAGGGGGGAGCAGTCTTAAATAATTTGTGAT 300  
Db 156 AspLeuLysProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAsp 175  
QY 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGAGTGTCTGTGG 360  
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195  
QY 361 ATGGCACCTGAAGTTTTTGAAGGTAGTAATTCAGTGAAGGAGTGTCTGTGG 420  
Db 196 MetAlaProGluValPheGluGlySerAsnTrpSerGluLysCysAspValPheSerTrp 215  
QY 421 GGTATTATTCTTTGGGAAGTATACGCGTTCGGAACCCCTTTGATGAGATTGGTGGCCCA 480  
Db 216 GlyIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
QY 481 GCTTTCCGATCATGTGGCTGTTCATATGTGTACTCGACCACTGATATAAAATTTA 540  
Db 236 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeu 255  
QY 541 CCTAACCCATTGAGAGCTGATGACTCGTGTGCTTAAAGATCTTCCAGAGCCCT 600  
Db 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
QY 601 TCAATGGAGAAATTGTGAAATAATGACTCACTTGTGGTCTTAAAGATCTTCCAGAGCAGAT 660  
Db 276 SerMetGluGluValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
QY 661 GAGCCATTACAGTATCTTGTCTAG 684  
Db 296 GluProLeuGlnTyrProCysGln 303  
RESULT 3  
JC5956  
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human  
C:Species: Homo sapiens (man)  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: JC5956  
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.  
Biochem. Biophys. Res. Commun. 243, 545-549, 1998  
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind  
A:Reference number: JC5955; MUID:98153801; PMID:9480845  
A:Accession: JC5956  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-606 <SAK>  
A:Cross-references: DBJ:AB009357; NID:G2924625; PIDN:BAA25026.1; PID:G2924626  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C:Keywords: phosphotransferase  
Alignment Scores:  
Pred. No.: 3 9e-73 Length: 606  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0  
US-09-830-144-1\_copy\_408\_1091 (1-684) x JC5956 (1-606)  
QY 1 GTAGAGCTTCGGCAGTTATCCGTTGAAACCATCTTAATTTGTAAGCTTTATGAGCC 60  
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95

QY 61 TGCTGAATCCAGTGTGCTTCTGATGGAATATGCTGAAGGGGCTCTTTATATATCTG 120  
 Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115  
 QY 121 CTGCATGCTGCTGAACCATGCTCATATATATCTGCTGCCACGCAATCAGTTGGTGT 180  
 Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 135  
 QY 181 CAGTGTTCCTCCAGAGTGGCTTATCTTCACAGCATGCCAACCAAGCGCTAATTCACAGG 240  
 Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisarg 155  
 QY 241 GACCTGAACCAACCAACCTTACTGCTGTTGAGGGGACAGTCTTAAATTTGTGAT 300  
 Db 156 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysileCysAsp 175  
 QY 301 TTTGCTACAGCTGTGACATTCAGACACATGACCAATTAACAAGGGAGTGTCTGCTGG 360  
 Db 176 PheGlyThrAlaCysaspGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195  
 QY 361 ATGCACCTCGAAGTGTGTTTGAAGTAGTAATTAACAGTGAATAATGTGACCTTCAGCTGG 420  
 Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysaspValPheSerTyr 215  
 QY 421 GGTATTAATCTTTGGGAAGTATAACGCTGCGAAACCTTTTCATGAGATTCGTGGCCCA 480  
 Db 216 GlyIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235  
 QY 481 GCTTTCCGAATCATGGGCTGTTCATATGCTGACTCGACCACTGATGATAAAATTTA 540  
 Db 236 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProLeuLeuLysAsnLeu 255  
 QY 541 CCTAACCCCATTCAGAGCCCTGATGACTCGTGTGTTGCTTAAAGATCTTCCAGCGCCCT 600  
 Db 256 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 275  
 QY 601 TCAATGGAGAAATTTGAAATTAATGACTTCATGCTGCTGCTGCTGCTGCTGCTGCTG 660  
 Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
 QY 661 GAGCCATTACAGTATCTCTCTCAG 684  
 Db 296 GluProLeuGlnTyrProCysGln 303  
 RESULT 4  
 S68178  
 mixed-lineage protein kinase 2 (EC 2.7.1.-) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 11-Jun-1999  
 C:Accession: S68178; I38044; S32468  
 R:Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simps  
 Eur. J. Biochem. 234, 492-500, 1995  
 A:Title: Complete nucleotide sequence, expression, and chromosomal localisation of human  
 A:Reference number: S68178; MUID:96128179; PMID:8536694  
 A:Accession: S68178  
 A:Molecule type: mRNA  
 A:Residues: 1-954 <DOR>  
 A:Cross-references: EMBL:X90846; NID:9971419; PIDN:CAA62351.1; PID:9971420  
 R:Kato, M.; Hirai, M.; Sugimura, T.; Terada, M.  
 Oncogene 10, 1447-1451, 1995  
 A:Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase  
 A:Reference number: I38044; MUID:95249256; PMID:7731697  
 A:Accession: I38044  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-461, 'A', 'V', 465-470, 'S', 472-806, 'R', 808-817, 'A', 819-954 <RES>  
 A:Cross-references: EMBL:Z48615; NID:9758592; PIDN:CAA88531.1; PID:9758593  
 R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.  
 Eur. J. Biochem. 213, 701-710, 1993  
 A:Title: Identification of a new family of human epithelial protein kinases containing b  
 A:Reference number: S32467; MUID:93238756; PMID:8477742  
 A:Accession: S32468  
 A:Molecule type: mRNA

A:Residues: 244-464, 'AQAAGRRQPHQPALWL' <DO2>  
 C:Genetics:  
 A:Gene: GDB:MLK2; GDB:MST  
 A:Cross-references: GDB:362654; GDB:624810; OMIM:600137  
 A:Map position: 19q13.1-19q13.2  
 C:Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology  
 C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein ki  
 F:23-76/Domain: SH3 homology <SH3>  
 F:96-364/Domain: protein kinase homology <KIN>  
 F:104-112/Region: protein kinase ATP-binding motif  
 F:384-405/Region: leucine zipper motif  
 F:419-440/Region: leucine zipper motif  
 F:449-463/Region: basic  
 F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Alignment Scores: 3.94e-18 Length: 954  
 Pred. No.: 392.50 Matches: 88  
 Score: 56.68% Conservative: 35  
 Percent Similarity: 56.68% Mismatches: 77  
 Best Local Similarity: 40.55% Indels: 17  
 Query Match: 31.35% Gaps: 6  
 DB: 1

US-09-830-144-1\_COPY\_408\_1091 (1-684) x S68178 (1-954)

QY 4 GAGCTCGCAGTATCCCGTGTGAACCATCCTAATATTGTAAAGCTTTATGAGCCTGC 63  
 Db 145 GluAlaArgLeuPheGlyAlaLeuGlnHisProAsnIleAlaLeuArgGlyAlaCys 164  
 QY 64 TTGAATCCCA-----GTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAAT 117  
 Db 165 LeuAsnProHisLeuCysLeuValMetGluTyrAlaArgGlyGlyAlaLeuSerArg 184  
 QY 118 GTGCTCATGCTGCTGAACCATGCTCATATATTACTGCTGCCACGCAATGAGTGTGT 177  
 Db 185 ValLeuAlaGlyArgValProHisVal-----LeuValAsnTyrPala 200  
 QY 178 TTACAGTGTTCCTCAAGAGTGGCTTATCTTCAGCATGCAACCCAAAGCGCTAATTCAC 237  
 Db 201 ValGlnValAlaArgGlyMetAsnTyrLeuHisAsnAspAlaProValProIleIleHis 220  
 QY 238 AGGAGCTGAACACCAACCACTTACTGCTGGT-----GCGGG 276  
 Db 221 ArgAspLeuLysSerIleAsnIleLeuLeuAlaIleGluAsnHisAsnLeuAla 240  
 QY 277 GGCACAGTCTCAAAATTTGTGATTTTGTGACAGCTGTGAC---ATTGACACACATG 333  
 Db 241 AspThrValLeuLysIleThrAspPheGlyLeuAlaArgGluTyrHisLysThrThrLys 260  
 QY 334 ACCAATAACAAGGGAGTGTGCTTGGATGCGACCTGAAGTTTGAAGGTAGTAATTCAC 393  
 Db 261 MetSerAlaAlaGlyThrTyrAlaTyrMetAlaProGluValIleArgLeuSerLeuPhe 280  
 QY 394 AGTGAAAAATGTGACGCTTTGAGTGGGGTATTATCTTTGGGAAGTGAATACCGCTCGG 453  
 Db 281 SerLysSerSerAspValTyrSerPheGlyValLeuLeuTyrGluLeuThrGlyGlu 300  
 QY 454 AACCCCTTTGATGATGTTGGTGGCCAGCTTCCGATCATGTGGGCTGT---CATAAT 510  
 Db 301 ValProTyrArgGluIle-----AspAlaLeuAlaValAlaTyrGlyValAlaMetAsn 318  
 QY 511 GGTACTCGACCAACCAACCAATTTTACCTAAGCCCATTTGAGCGCTGATGACTCGT 570  
 Db 319 LysLeuThrLeuProIleProSerThrCysProGluProPheAlaArgLeuLeuGluGlu 338  
 QY 571 TGTGGTCTAAAGATCTTCCAGCGCCCTTCAATGGAGGAAATGTGAAA 621  
 Db 339 CysTyrAspProAspProHisGlyArgProAspPheGlySerIleLeuLys 355

RESULT 5  
 S684635  
 probable protein kinase [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)



RESULT 7  
JU0229  
mixed-lineage protein kinase 1 - human  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999  
C/Accession: S32467; JU0229  
R;Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.  
Eur. J. Biochem. 213, 701-710, 1993  
A:Title: Identification of a new family of human epithelial protein kinases containing a  
A:Reference number: S32467; MUID:93238756; PMID:8477742  
A:Accession: S32467  
A:Molecule type: mRNA  
A:Residues: 1-394 <DO2>  
C:Genetics:  
A:Gene: GDB:MLK1  
A:Cross-references: GDB:141921; OMIM:600136  
A:Map position: 14Q24.3-14Q31  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k  
F;1-269/Domain: protein kinase homology <KIN>  
F;1-269/Domain: catalytic <CAT>  
F;9-17/Region: protein kinase ATP-binding motif  
F;289-310/Region: leucine zipper motif  
F;324-345/Region: leucine zipper motif  
F;354-368/Region: basic

Alignment Scores:  
Pred. No.: 3,49e-16 Length: 394  
Score: 363.50 Matches: 83  
Percent Similarity: 54.09% Conservative: 36  
Best Local Similarity: 37.73% Mismatches: 84  
Query Match: 29.03% Indels: 17  
DB: 2 Gaps: 6

US-09-830-144-1\_COPY\_408\_1091 (1-684) x JU0229 (1-394)

QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGAGCCTGC 63  
Db 50 GluAlaLysLeuPheAlaMetLeuLysHisProAsnIleAlaLeuArgGlyValCys 69  
QY 64 TTGAATCCA-----GTGTGCTGTGATGGAATATGCTGAAGGGGCTCTTTATATAAT 117  
Db 70 LeuLysGluProAsnLeuCysLeuValMetGluPheAlaArgGlyGlyProLeuAsnArg 89  
QY 118 GTGCTGCATGCTGTGAACCATCTTATATATCTGCTGCCACGCAATGAGTTGGTGT 177  
Db 90 ValLeuSerGlyLysArgGlyProProAspIle-----LeuValAsnTIPAla 105  
QY 178 TTACAGTGTCCCAAGAGTGGCTTATCTTACAGCATGCAACCAAGCGCTAATTCAC 237  
Db 106 ValGlnIleAlaArgGlyMetAsnTyrLeuHisAspGluAlaIleValProIleIleHis 125  
QY 238 AGGACCTGAAACCAACCAACTTACTGCTGTT-----GCAGGG 276  
Db 126 ArgAspLeuLysSerSerAsnIleLeuLeuGlnLysValGluAsnGlyAspLeuSer 145  
QY 277 GGGACAGTCTTAAATAATTTGATTTTGGTACGCTGTGAC---ATTACAGACACACATG 333  
Db 146 AsnLysIleLeuLysIleThrAspPheGlyLeuAlaArgGluTyrHisArgThrThrllys 165  
QY 334 ACCAATAACAAGGAGTGGCTTGTGGATGGCAGCTGAAGTTTGAAGGTAGTAATAC 393  
Db 166 MetSerAlaAlaGlyThrTyraLalrPMetAlaProGluValIleArgAlaSerMetPhe 185  
QY 394 AGTGAAAAATGTAGCTCTTACGTGGGTATTTCTTTGGAGTGAATACGCGTGG 453  
Db 186 SerLysGlySerAspValTyrSerTyrGlyValLeuLeuTyrpGluLeuLeuThrGlyGlu 205  
QY 454 AAACCTTTGATGAGATTGGTGGCCAGCTTCCCAATCATGTGGGCTGTT---CATAT 510  
Db 206 ValProPheArgGlyIleAspGly-----LeuArgValAlaTyrGlyValAlaMetAsn 223

QY 511 GGTACTCGACCACTGATATAAAATTTTACCTAGCCATTGAGAGCCTGTGACTCGT 570  
Db 224 LysLeuAlaLeuProIleProSerThrCysProGluProPheAlaLysLeuMetGluAsp 243  
QY 571 TGTGTGCTTAAGATATCTCCAGCCCTTCAATGGAGAAATTTGTAATAATGACT 630  
Db 244 CysTIPAsnProAspProHisSerArgProSerPheThrAsnIleLeuAspGlnLeuThr 263

RESULT 8  
T18287  
protein-tyrosine kinase (EC 2.7.1.112) - slime mold (Dictyostelium discoideum)  
C:Species: Dictyostelium discoideum  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 04-Feb-2000  
C/Accession: T18287  
R;Adler, K.  
submitted to the EMBL Data Library, July 1996  
A:Reference number: Z18856  
A:Accession: T18287  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1338 <ADL>  
A:Cross-references: EMBL:U64830; NID:G1468982; PID:G1468983; PIDN:AB04999.1  
C:Genetics:  
A:Introns: 1181/3  
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Alignment Scores:  
Pred. No.: 3,27e-16 Length: 1338  
Score: 362.00 Matches: 80  
Percent Similarity: 57.62% Conservative: 41  
Best Local Similarity: 38.10% Mismatches: 73  
Query Match: 28.91% Indels: 16  
DB: 2 Gaps: 6

US-09-830-144-1\_COPY\_408\_1091 (1-684) x T18287 (1-1338)

QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGAGCCTGC 63  
Db 1100 GluValSerSerLeuIleLysSerHisProAsnValValThrPheMetGlyAlaArg 1119  
QY 64 TTGAATCCAAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGCTG 123  
Db 1120 IleAspProCysIlePheThrGluTyrLeuGlnGlyGlySerLeuTyrAspValLeu 1139  
QY 124 CAT-----GGTGTGAACCATTCCTATATATCTGCTGCCACGCAATGAGT 171  
Db 1140 HisIleGlnLysIleLysLeuAsnProLeuMetMetTyrLysMetIleHisAspLeu--- 1158  
QY 172 TGGTGTTTACAGTGTCCCAAGGAGTGGCTTATCTTACAGCATGCAACCAAGCGCTA 231  
Db 1159 -----SerLeuGlyMetGluHisLeuHisSerIleGln-----Met 1170  
QY 232 ATTCACAGGACCTGAAACCAACCACTTACTGCTGTGTCAGGGGGGACAGTCTTAAAA 291  
Db 1171 LeuHisArgAspLeuThrSerLysAsnIleLeuLeuAspGluPheLysAsnIle---Lys 1189  
QY 292 ATTTGTGATTTGTGTACAGCTGTGACATTCAGACACACATGACC---AATAACAAGGG 348  
Db 1190 IleAlaAspPheGlyLeuAlaThrLeuSerAspAspMetThrLeuSerGlyIleThr 1209  
QY 349 AGTCTGCTTGGATGGCACCTGAAAGTTTTCGAAGGTAGTAATACAGTGAATAATGTGAC 408  
Db 1210 AsnProArgTIPArgSerProGluLeuThrLysGlyLeuValTyrAsnGluLysValAsp 1229  
QY 409 GTCTTCAGCTGGGTATTTCTTTGGAGTGTATAACCGCTCGGAACCCCTTTGATGAG 468  
Db 1230 ValTyrSerPheGlyLeuValValTyrGluIleTyrThrGlyLysIleProPheGluGly 1249  
QY 469 ATTGGTGGCCAGCTTCCGGAATCATGTGGCTGTTCTATATGTTACTCGACCCACCTG 528  
Db 1250 LeuAspGlyThrAlaSerAlaAlaLysAlaPheGluAsn---TyrArgProAlaIle 1268  
QY 529 ATAAAAATTTTACCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTTGTTCTTAAAGATCCT 588

```
Db 1269 ProProAspCysProValSerLeuArgLysLeuIleThrLysCysTrpAlaSerAspPro 1288
    ::: ||| ::: |||:::||||:::||||::: |||||
QY TCCACAGCGCCCTCAATGGAGGAATGTG 618
    ||||| ||||| ||||| ||||| |||||
Db 1289 SerGlnArgProSerPheThrGluIleLeu 1298
    ||||| ||||| ||||| ||||| |||||
RESULT 9
S29851
protein kinase 6 (EC 2.7.1.-) - soybean
C:Species: Glycine max (soybean)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C:Accession: S29851; S27760
R:Feng, X.H.; Zhao, Y.; Bottino, P.J.; Kung, S.
Biochim. Biophys. Acta 1172, 200-204, 1993
A:Title: Cloning and characterization of a novel member of protein kinase family from soybean
A:Reference number: S29851; MUID:931176812; PMID:8439562
A:Accession: S29851
A:Molecule type: mRNA
A:Residues: 1-462 <FEN>
A:Cross-references: EMBL:M67449; NID:g170046; PIDN:AAA34002.1; PID:g170047
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:154-419/Domain: protein kinase homology <KIN>
Alignment Scores:
Pred. No.: 1,02e-15 Length: 462
Score: 356.00 Matches: 76
Percent Similarity: 58.88% Conservative: 50
Best Local Similarity: 35.51% Mismatches: 74
Query Match: 28.43% Indels: 14
DB: 2 Gaps: 7
US-09-830-144-1_COPY_408_1091 (1-684) x S29851 (1-462)
QY 4 GAGCTTCGGCAGTATCCCGTGTGAACCATCTATATTTGTAAGCTTTATGAGCGCTGC 63
    |||:::||||:::||||:::||||::: |||||
Db 207 GluValThrLeuLeuSerArgLeuHisGlnAsnValIleLysPheSerAlaAlaCys 226
    |||:::||||:::||||:::||||::: |||||
QY 64 TTGAATCCA-----GTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTATATAT 117
    |||:::||||:::||||:::||||::: |||||
Db 227 ArgLysProValTyrCysIleIleThrGluTyrLeuAlaGluLysLeuArgAla 246
    |||:::||||:::||||:::||||::: |||||
QY 118 GTGTGTCATGTGCTGAACCATTCATATATATCTGCTGCC-----CACGCAATGAGT 171
    |||:::||||:::||||:::||||::: |||||
Db 247 TyrLeuHis-----LysLeuGluHisGlnThrIleSerLeuGlnLysLeuAla 263
    |||:::||||:::||||:::||||::: |||||
QY 172 TGGTCTTTACAGTGTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTA 231
    |||:::||||:::||||:::||||::: |||||
Db 264 PheAlaLeuAspIleAlaArgGlyMetGluTyrIleHisSer-----GlnGlyVal 280
    |||:::||||:::||||:::||||::: |||||
QY 232 ATTCAAGGACCTGAACCAACCAACTTACTGCTGTTGTCAGGGGGGACAGTCTTAAAA 291
    |||:::||||:::||||:::||||::: |||||
Db 281 IleHisArgAspLeuLysProGluAsn---IleLeuIleAsnGluAspHisLeuLys 299
    |||:::||||:::||||:::||||::: |||||
QY 292 ATTTGTGATTTGGTACAGCTGTGCATTCAGACA-----CACATGACCAATAACAG 345
    |||:::||||:::||||:::||||::: |||||
Db 300 IleAlaAspPheGlyIleAlaCysGluGluAlaSerCysAspLeuLeuAlaAspAspPro 319
    |||:::||||:::||||:::||||::: |||||
QY 346 GGGAGTGTCTGTGATGACCTGAAGTTTGAAGTAGTAAATTACAGTGAATAATGT 405
    |||:::||||:::||||:::||||::: |||||
Db 320 GlyThrTyrArgTrpMetAlaProGluMetIleLysArgLysSerTyrGlyLysLysVal 339
    |||:::||||:::||||:::||||::: |||||
QY 406 GACGCTTTGAGTGGGGTATTATCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTCAT 465
    |||:::||||:::||||:::||||::: |||||
Db 340 AspValTyrSerPheGlyLeuIleLeuTrpGluMetLeuThrGlyThrIleProTyrGlu 359
    |||:::||||:::||||:::||||::: |||||
QY 466 GAGATTGTGGCCAGCTTTCGGAATCATGTGGGCTGTTCATATGTTACTCGACCAACA 525
    |||:::||||:::||||:::||||::: |||||
Db 360 AspMet---AsnProIleGlnAlaAlaPheAlaValValAsnLysAsnSerArgProIle 378
    |||:::||||:::||||:::||||::: |||||
QY 526 CTGATAAAAAATTTACCTAAGCCATTGAGAGCTGATGACTGCTGTGCTTAAAGAT 585
    |||:::||||:::||||:::||||::: |||||
```

```
Db 379 IleProSerAsnCysProProAlaMetArgAlaLeuIleGluInCysTrpSerLeuGln 398
    |||:::||||:::||||:::||||::: |||||
QY 586 CTTTCCAGCGCCCTTCAATGGAGGAATGTGAAATAATG 627
    |||:::||||:::||||:::||||::: |||||
Db 399 ProAspLysArgProGluPheTrpGlnValValLysIleLeu 412
    |||:::||||:::||||:::||||::: |||||
RESULT 10
T46150
protein kinase ATN1-like protein - Arabidopsis thaliana
N:Alternate names: protein T3A5.110
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000
C:Accession: T46150; T08394
R:Blöcker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Quétier, F.; Sala
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z23024
A:Accession: T46150
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-370 <BLO>
A:Cross-references: EMBL:AL132979
A:Experimental source: cultivar Columbia; BAC clone T3A5
R:Quétier, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salanoub
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16409
A:Accession: T08394
A:Molecule type: DNA
A:Residues: 1-370 <BLO>
A:Cross-references: EMBL:AL049862; GSPDB:GN00061; ATSP:F18B3.10
A:Experimental source: cultivar Columbia; BAC clone F18B3
C:Genetics:
A:Gene: ATSP:F18B3.10
A:Map position: 3
A:Introns: 54/2; 99/3; 164/1; 204/3; 257/3
A:Note: T3A5.110
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
Alignment Scores:
Pred. No.: 1.15e-15 Length: 370
Score: 355.50 Matches: 79
Percent Similarity: 56.05% Conservative: 46
Best Local Similarity: 35.43% Mismatches: 81
Query Match: 28.39% Indels: 17
DB: 2 Gaps: 8
US-09-830-144-1_COPY_408_1091 (1-684) x T46150 (1-370)
QY 4 GAGCTTCGGCAGTATATCCCGTGTGAACCATCTATATTTGTAAGCTTTATGAGCGCTGC 63
    |||:::||||:::||||:::||||::: |||||
Db 85 GluValLeuLeuLeuSerLysMetLysHisAspAsnIleValLysPheValGlyAlaCys 104
    |||:::||||:::||||:::||||::: |||||
QY 64 TTGAATCCA---GTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTATATATATG 120
    |||:::||||:::||||:::||||::: |||||
Db 105 IleGluProGlnLeuIleValThrGluLeuValGluGlyGlyThrLeuGlnArgPhe 124
    |||:::||||:::||||:::||||::: |||||
QY 121 CTGATGTGTGTGACCATTCGCATATATATCTGCTGCCAGCAATAGATTGGTGT 180
    |||:::||||:::||||:::||||::: |||||
Db 125 Methis---SerArgProGlyPro---LeuAspLeuLysMetSerLeuSerPheAlaLeu 142
    |||:::||||:::||||:::||||::: |||||
QY 181 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 240
    |||:::||||:::||||:::||||::: |||||
Db 143 AspIleSerArgAlaMetGluPheValHisSer-----AsnGlyIleIleHisArg 159
    |||:::||||:::||||:::||||::: |||||
QY 241 GACCTGAAACCAACCAACTTACTGCTGTTGTCAGGGGGACAGTCTCTAAAATTGTGAT 300
    |||:::||||:::||||:::||||::: |||||
Db 160 AspLeuAsnProArgAsnLeuLeuValThrGlyAspLeuLysHisValLysLeuAlaAsp 179
    |||:::||||:::||||:::||||::: |||||
QY 301 TTTGGTACAGCTCT---GACATTACAGACACATGACCAATACACAGGGGAGTCTCTCT 357
    |||:::||||:::||||:::||||::: |||||
Db 180 PheGlyIleAlaArgGluThrArgGlyGlyMetThrCysGluAlaGlyThrSerLys 199
    |||:::||||:::||||:::||||::: |||||
QY 358 TGGATGCGACCTGAAGTTT---GAAGGTAGTAATTAC 393
    |||:::||||:::||||:::||||::: |||||
```

Db 200 TrpMetAlaProGluValTyrSerProGluProLeuArgValGlyGluLysGluTyr 219  
 QY 394 AGTGAATAATGTGACGCTCTTACGCTGGGTATTATTCTTTTGGGAAGTGAATAACGCTCGG 453  
 Db 220 AspHisLysAlaAspIleTyrSerPheAlaIleValLeuTrpGlnLeuValThrAsnGlu 239  
 QY 454 AAACCTTTGATGAGATTGGTGGCCAGCTTTCCGAATCATGTGGCTGTTTCATAATGGT 513  
 Db 240 GluProPheProAspValProAsnSerLeuPhe---ValProTyrLeuValSerGlnGly 258  
 QY 514 ACTCGACCCACTGATGATAAAATTTTACCTTAAGCCATTGAGAGCTGTGATGACTCGTGT 573  
 Db 259 ArgArgProIleLeuThrIysThr---ProAspValPheValProIleValGluSerCys 277  
 QY 574 TGTCTCTAAAGATCTTCCAGCCGCTTCAATCGAGGAATTTGAAATAATATGACTCAC 633  
 Db 278 TrpAlaGlnAspProAspAlaArgProGluPheLysGluIleSerValMetLeuThrAsn 297  
 QY 634 TTGATCGGG 642  
 Db 298 LeuLeuArg 300

RESULT 11  
 JC2363  
 protein kinase (EC 2.7.1.37) ZPK - human  
 N;Alternate names: leucine-zipper protein kinase  
 C;Species: Homo sapiens (man)  
 C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Feb-2001  
 C;Accession: JC2363  
 R;Reddy, U.R.; Pleasure, D.  
 Biochem. Biophys. Res. Commun. 202, 613-620, 1994  
 A;Title: Cloning of a novel putative protein kinase having a leucine zipper domain from  
 A;Reference number: JC2363; MUID:94311945; PMID:8037767  
 A;Accession: JC2363  
 A;Molecule type: mRNA  
 A;Residues: 1-668 <RED>  
 A;Cross-references: EMBL:U07358  
 A;Experimental source: Brain  
 A;Note: The nucleotide sequence for this amino acid sequence is inconsistent with that f  
 he codon ACC for residue 661 as Pro, the codon GAACACCTCTCCA for residues 664-668 as A  
 C;Comment: This protein belongs to the family of non-receptor kinase.  
 C;Genetics:  
 A;Gene: GDB:ZPK  
 A;Cross-references: GDB:383963; OMIM:600447  
 A;Map position: 12q13-12q13  
 C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 C;Keywords: Atp; leucine zipper; nucleotide binding; P-loop; phosphotransferase  
 F;123-137/Domain: protein kinase homology <KIN>  
 F;131-139/Region: protein kinase ATP-binding motif  
 F;443-471/Region: leucine zipper motif  
 F;538-545/Region: nucleotide-binding motif A (P-loop)  
 F;152/Active site: Lys #status predicted

Alignment Scores:  
 Pred. No.: 1.35e-15 Length: 668  
 Score: 353.50 Matches: 75  
 Percent Similarity: 54.63% Conservative: 43  
 Best Local Similarity: 34.72% Mismatches: 81  
 Query Match: 28.23% Indels: 17  
 DB: 2 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x JC2363 (1-668)

QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCCTGC 63  
 Db 161 AspileLysHisLeuArgLysLeuLysHisProAsnIleIleThrPheLysGlyValCys 180  
 QY 64 TTGAATCCA-----GTGTGTTCTGTGATGGAATATGTGAAGGGGCTCTTTATATAT 117  
 Db 181 ThrGlnAlaProCysTyrCysIleLeuMetGluPheCysAlaGlnGlnLeuTyrGlu 200  
 QY 118 GTGCTGCATGGTGTGAACCATTCGCATATATATCTGCTGCCACGCAATGATGGTGT 177

Db 201 ValLeuArgAlaGlyArgProVal-----ThrProSerLeuLeuValAspTrpSer 217  
 QY 178 TTACAGTGTTCCTCAAGGAGTGCCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCC 237  
 Db 218 MetGlyIleAlaGlyGlyMetAsnTyrLeuHisLeuHisLys-----IleIleHis 234  
 QY 238 AGGAGCTCAAAACCCAACTTACTGCTGGTTCAGGGGGGACACAGTCTCTAAATAATTTGT 297  
 Db 235 ArgAspLeuLysSerProAsn---MetLeuIleThrTyrAspValValLysIleSer 253  
 QY 298 GATTTTGGTACGCCCTGTGACATTCAGACACATGACCAATAACAAG-----GGAGT 351  
 Db 254 AspPheGlyThrSerLysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThr 273  
 QY 352 GCTGCTGCATGCACCTGAAGTTTGAAGGTAGTAAATTACAGTGAATAATGTGAGTCA 411  
 Db 274 ValAlaTrpMetAlaProGluValIleArgAsnGluProValSerGluLysValAspIle 293  
 QY 412 TTCAGCTGGGGTATTATTCTTTGGGAAGTGAATACCGTCGGAAACCCCTTTGATGAGATT 471  
 Db 294 TrpSerPheGlyValValLeuTrpGluLeuLeuThrGlyGluIleProTyrLysAspVal 313  
 QY 472 GGTGGCCAGCTTCCGAATCATGTGGGTGTT---CATTAATGGTACTCCACACCACTG 528  
 Db 314 AspSerSerAla-----IleIleTrpGlyValGlySerAsnSerLeuHisLeuProVal 331  
 QY 529 ATAAAAATTTACCTTAAGCCATTGAGAGCTGATGACTGTTGTTGGTCTTAAAGATCCT 588  
 Db 332 ProSerSerCysProAspGlyPheLysIleLeuLeuArgGlnCysTrpAsnSerLysPro 351  
 QY 589 TCCAGCGCCCTTCAATGGAGGAATTTGAAATAATATGACTCACTCTG 636  
 Db 352 ArgAsnArgProSerPheArgGln-----IleLeuLeuHisLeu 364

RESULT 12  
 A55318  
 serine/threonine protein kinase (EC 2.7.1.1-) DLK - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 24-Sep-1999  
 C;Accession: A55318  
 R;Holzman, L.B.; Merritt, S.E.; Fan, G.  
 J. Biol. Chem. 269, 30808-30817, 1994  
 A;Title: Identification, molecular cloning, and characterization of dual leucine zipper  
 S;Reference number: A55318; MUID:95074107; PMID:7983011  
 A;Accession: A55318  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-888 <HOL>  
 A;Cross-references: GB:U14636; NID:G602677; PIDN:AAA57280.1; PID:G602678  
 C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
 C;Keywords: Atp; leucine zipper; phosphotransferase  
 F;156-404/Domain: protein kinase homology <KIN>  
 F;164-172/Region: protein kinase ATP-binding motif

Alignment Scores:  
 Pred. No.: 1.26e-15 Length: 888  
 Score: 353.50 Matches: 75  
 Percent Similarity: 54.63% Conservative: 43  
 Best Local Similarity: 34.72% Mismatches: 81  
 Query Match: 28.23% Indels: 17  
 DB: 2 Gaps: 8

US-09-830-144-1\_COPY\_408\_1091 (1-684) x A55318 (1-888)

QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCCTGC 63  
 Db 194 AspileLysHisLeuArgLysLeuLysHisProAsnIleIleThrPheLysGlyValCys 213  
 QY 64 TTGAATCCA-----GTGTGTTCTGTGATGGAATATGTGAAGGGGCTCTTTATATAT 117  
 Db 214 ThrGlnAlaProCysTyrCysIleLeuMetGluPheCysAlaGlnGlnLeuTyrGlu 233





Search completed: December 4, 2003, 09:08:38



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 07:06:54 ; Search time 16.1757 Seconds  
(without alignments)  
3977.116 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091

Perfect score: 1252  
Sequence: 1 gtagagcttcgcagttatc.....cattacagttatctgttcag 684

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09830144/runat\_03122003\_122344\_21226/app\_query.fasta\_1.1230  
-DB=SwissProt\_41 -QFMT=fastan -SURFIX=rsb -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=spt -THR\_MIN=100 -THR\_MAX=100 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 @CGN 1 1 16 @runat\_03122003\_122344\_21226 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	579	1 M3K7 MOUSE	Q62073 mus musculus
2	1252	100.0	606	1 M3K7 HUMAN	O43318 homo sapien
3	435.5	34.8	393	1 M3K7 DROME	P83104 drosophila
4	392.5	31.3	954	1 M3K4 HUMAN	Q02779 homo sapien
5	363.5	29.0	394	1 M3K9 HUMAN	P80192 homo sapien
6	355.5	28.4	888	1 M3K3 RAT	Q63796 rattus norv
7	353.5	28.2	859	1 M3K3 HUMAN	Q12852 homo sapien
8	353.5	28.2	888	1 M3K3 MOUSE	Q60700 mus musculus
9	341	27.2	410	1 KTK2 DICDI	P18161 dictyosteli
10	323	25.8	821	1 CTR1 ARATH	Q05609 arabidopsis
11	307.5	24.6	1115	1 RET MOUSE	P35546 mus musculus
12	304.5	24.3	1114	1 RET HUMAN	P07949 homo sapien
13	301.5	24.1	630	1 TEC MOUSE	P24604 mus musculus
14	301.5	24.1	631	1 TEC HUMAN	P42680 homo sapien
15	299.5	23.9	625	1 ITK MOUSE	Q03526 mus musculus
16	298.5	23.8	620	1 ITK HUMAN	HSSP; P08631; IAD5
17	294.5	23.5	2347	1 KROS HUMAN	MGD; MGI:1346877; Map3k7.
18	292.5	23.4	402	1 KROS_AVISU	InterPro; IPR000719; Prot_kinase.

19	292	23.3	507	1 KROS CHICK	P08941 gallus gall
20	290	23.2	1584	1 KYK1 DICDI	P18160 dictyosteli
21	288.5	23.0	628	1 TESI RAT	O63572 rattus norv
22	288	23.0	1307	1 PHYT_CERPU	P25848 ceratodon p
23	286	22.8	804	1 FPS_DROME	P18106 drosophila
24	284.5	22.7	832	1 ANR3 HUMAN	P57078 homo sapien
25	283.5	22.6	527	1 TKX MOUSE	P42682 mus musculus
26	280	22.4	822	1 FGR1 RAT	P41243 rattus norv
27	279	22.3	467	1 MATK RAT	Q04589 rattus norv
28	278	22.2	822	1 FGR1 HUMAN	P11362 homo sapien
29	278	22.2	822	1 FGR1 MOUSE	P16092 mus musculus
30	277.5	22.2	786	1 BTKL DROME	O08630 drosophila
31	275	22.0	984	1 EPA3 RAT	O08680 rattus norv
32	274.5	21.9	801	1 FGR3 MOUSE	O61851 mus musculus
33	271.5	21.7	1082	1 CC7 SCHPO	P41892 schizosacch
34	271	21.6	604	1 KRAA MOUSE	P04627 mus musculus
35	271	21.6	604	1 KRAA RAT	P14056 rattus norv
36	270.5	21.6	503	1 HCK RAT	P50545 rattus norv
37	270	21.6	806	1 FGR3 HUMAN	P22607 homo sapien
38	269.5	21.5	507	1 MATK HUMAN	P42679 homo sapien
39	269.5	21.5	626	1 TESI HUMAN	O15569 homo sapien
40	269	21.5	533	1 FPS_AVISP	P00541 avian sarco
41	269	21.5	819	1 FGR1 CHICK	P21804 gallus gall
42	269	21.5	983	1 EPA3 HUMAN	P29320 homo sapien
43	268	21.4	830	1 PKN2 MYXXA	P54736 myxococcus
44	268	21.4	1052	1 FAK1 HUMAN	O05397 homo sapien
45	268	21.4	1055	1 FAK1 RAT	O35346 rattus norv

## ALIGNMENTS

RESULT 1					
M3K7 MOUSE					
ID M3K7 MOUSE	STANDARD;	PRT;	579 AA.		
AC Q62073;					
DT 16-OCT-2001 (Rel. 40, Last sequence update)					
DT 16-OCT-2001 (Rel. 40, Last annotation update)					
DE Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)					
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).					
GN MAP3K7 OR TAK1.					
OS Mus musculus (Mouse).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX NCBI_TaxID=10090;					
RN [1]					
RP SEQUENCE FROM N.A.					
EX MEDLINE=96123277; PubMed=8533096;					
RA Yanaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,					
RA Taniguchi T., Nishida E., Matsumoto K.;					
RT Identification of a member of the MAPKKK family as a potential					
RT mediator of TGF-beta signal transduction."					
RL Science 270:2008-2011(1995).					
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATES YET UNDEFINED MAPKKS.					
CC -!- MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B					
CC ACTIVATION.					
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.					
CC MAP KINASE KINASE KINASE SUBFAMILY.					
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C210RF7.					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC the European Bioinformatics Institute. There are no restrictions on its					
CC use by non-profit institutions as long as its content is in no way					
CC modified and this statement is not removed. Usage by and for commercial					
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC or send an email to license@isb-sib.ch).					
CC -----					
CC EMBL; D76446; BAA11184.1; -					
CC HSSP; P08631; IAD5.					
CC MGD; MGI:1346877; Map3k7.					
CC InterPro; IPR000719; Prot_kinase.					

DR InterPro; IPR002290; Ser thr pkinase.  
 DR InterPro; IPR001245; Tyr pkinase.  
 DR Pfam; PF00069; pkinase.1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
 FT DOMAIN 8 16 POLY-SER.  
 FT NP\_BIND 36 291 PROTEIN KINASE.  
 FT NP\_BIND 42 50 ATP (BY SIMILARITY).  
 FT BINDING 63 63 ATP (BY SIMILARITY).  
 FT ACT\_SITE 156 156 BY SIMILARITY.  
 SQ SEQUENCE 579 AA; 64227 MW; 97C8F6F3C8E283EE CRC64;

Alignment Scores:  
 Pred. No.: 2,82e-116 Length: 579  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 1 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x M3K7\_MOUSE (1-579)

QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCCTTAATATTGTAAGCTTTATGAGCC 60  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnileValLysLeuTyrGlyAla 95  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 61 TGCTTGAATCCAGTGTCTTGTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATGTC 120  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 121 CTCATGGTCTGAACCATTCATATATCTGCTGCCACCAATGATGTGCTGTTA 180  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 135  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 181 CAGTGTTCACAGAGTGGCTTCTTCACAGATCCACCAACCAAGCGCTAATTCACAGG 240  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArg 155  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 241 GACCTGAACACCAACTTACTGCTGCTGCTGAGGGGACAGCTTCTAAATAATTTGTGAT 300  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 156 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysLleCysAsp 175  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 301 TTGTGTACAGCTGTGACATTCAGACACATGACCAATTAACAGGGAGTGCTGCTTG 360  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 361 ATGGCACCTGAAGTCTTTGAAGTACTAATACAGTGAATAATGTACGCTCTTCAGCTGG 420  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 215  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 421 GSTATTATTCTTTGGGAAGTGATAGCGCTCGGAACCCCTTTGATCAGATTTGTCGCCCA 480  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 216 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 235  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 481 GCTTTCGGAATCATGTGGCTGTTTCAATATGTTACTGACACCACTCATGATAAAATTTA 540  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 236 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProLeuIleLysAsnLeu 255  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 541 CTTAAGCCCATTCAGAGCTGTGACTCGTGTGTTGTTGTTAAAGATCTTCCAGCGCCCT 600  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 256 ProLysProIleGluSerLeuMetThrArgCysTyrPserLysAspProSerGlnArgPro 275  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 601 TCAATGAGGAATTCGAAATAATAGTCACTGATCGGTACTTCTCCAGAGACAGAT 660  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 661 GAGCCATTCAGTATCTTGTCTAG 684  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 296 GluProLeuGlnTyrProCysGln 303  
 Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 2  
 M3K7 HUMAN STANDARD; PRT; 606 AA.  
 ID O43318; O43317; O43319;  
 AC O43318; O43317; O43319;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)  
 DE (transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).  
 DE MAP3K7 OR TAK1.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1A; 1B AND 1C).  
 RC TISSUE=lung;  
 RX MEDLINE=98153801; PubMed=9480845;  
 RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;  
 RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism";  
 RL Biochem. Biophys. Res. Commun. 243:545-549(1998).  
 CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS. MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B ACTIVATION.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1B;  
 CC IsoId=O43318-1; Sequence=Displayed;  
 CC Name=1A;  
 CC IsoId=O43318-2; Sequence=VSP\_004886;  
 CC Name=1C;  
 CC IsoId=O43318-3; Sequence=VSP\_004887. VSP\_004888;  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP KINASE KINASE KINASE SUBFAMILY.  
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AB0093157; BAA25026.1; -;  
 CC EMBL; AB0093156; BAA25025.1; -;  
 CC EMBL; AB0093158; BAA25027.2; -;  
 CC PIR; JC5955; JC5955.  
 CC PIR; JC5956; JC5956.  
 CC HSP; P08631; IAD5.  
 CC Genew; HGNC:6859; MAP3K7.  
 CC MIM; 602614;  
 CC GO; GO:0004709; F:MAP kinase kinase activity; TAS.  
 CC GO; GO:0007179; P:TGFbeta receptor signaling pathway; TAS.  
 CC InterPro; IPR000719; Prot kinase.  
 CC InterPro; IPR002290; Ser thr pkinase.  
 CC InterPro; IPR001245; Tyr pkinase.  
 CC Pfam; PF00069; pkinase; 1.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC ProDom; PD000001; Prot kinase; 1.  
 CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 CC PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 CC PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;  
 KW Alternative splicing.  
 FT DOMAIN 8 14 POLY-SER.  
 FT DOMAIN 36 291 PROTEIN KINASE.  
 FT NP\_BIND 42 50 ATP (BY SIMILARITY).  
 FT BINDING 63 63 ATP (BY SIMILARITY).  
 FT ACT\_SITE 156 156 BY SIMILARITY.

```

FT VARSPLIC 404 430 Missing (in isoform 1A).
FT FTID=VSP_004886.
FT VARSPLIC 509 518 PLAPCNSKE -> ARTSRTGPG (in isoform 1C).
FT FTID=VSP_004887.
FT VARSPLIC 519 606 Missing (in isoform 1C).
FT FTID=VSP_004888.
FT SEQUENCE 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;

Alignment Scores:
Pred. No.: 2,84e-116 Length: 606
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-830-144-1_COPY_408_1091 (1-684) x M3K7_HUMAN (1-606)

QY 1 GTGAGCTTCGGAGTATCCCGTGAACCATCTTAATTCGTAAGCTTTATGGAGCC 60
Db 76 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyAla 95
QY 61 TGCTTGAATCCAGTGTCTGTGTGATGCAATATGCTGAAGGGGCTCTTTATATAATGTG 120
Db 96 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 115
QY 121 CTGATGCTGCTGAACCATTCGCAATATTAATGCTGCTGCCACGCAATGAGTTGGTTTA 180
Db 116 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 135
QY 181 CAGTGTCCCAAGGAGTGGCTTATCTTCACAGATGCAACCCAAAGCGCTAATTCACAGG 240
Db 136 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 155
QY 241 GACCTGAACACCAAACTTACTGCTGCTGAGGGGAGCAGTTCTAAAAATTTGTGAT 300
Db 156 AspLeuLysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175
QY 301 TTGTTGACAGCTGTGACATTCAGACACATGACCAATACAGGGGAGTGTGCTGGTGG 360
Db 176 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 195
QY 361 ATGGCAGCTGAAGTTTGTGAGGTAGTAAATACAGTGAATAATGTCAGCTCTCAGCTGG 420
Db 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 215
QY 421 GGTATTATCTTGGGAGTGATACCGCTCGGAACCCCTTTGATGAGATTGGTGCCCA 480
Db 216 GlyIleLeuLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235
QY 481 GCTTTCGGAATCATGTGGCTGTTTCATATGTTACTCGACACCATCATGATAAAATTTA 540
Db 236 AlaPheArgIleMetTyrAlaValAlaHisAsnGlyThrArgProProLeuIleLysAsnLeu 255
QY 541 CTTAAGCCCATGAGAGCTGATCAGCTGTGTGGTCTTAAGATCTCTCCAGCGCCT 600
Db 256 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 275
QY 601 TCATGAGGAAATTTGMAAATAATGACTCAGCTGATGCGGTACTTTCCAGGAGCAGAT 660
Db 276 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 295
QY 661 GAGCATTACAGTATCTCTGTGTCAG 684
Db 296 GluProLeuGlnTyrProCysGln 303

RESULT 3
M3K7_DROME
ID M3K7_DROME STANDARD; PRT; 393 AA.
AC P81104;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE DE Putative mitogen-activated protein kinase kinase 7
DE (EC 2.7.1.-).
GN TAKL1.
OC Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.P.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.A.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foelel C., Gabriell A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RC CONCEPTUAL TRANSLATION.
RP Manning G., Sudarsanam S., Plowman G.;
RA "Prediction of novel protein kinases from the Drosophila genome
RT project and EST sequences.";
RL Unpublished observations (AUG-2001).
CC -!- FUNCTION: Can phosphorylate and activate yet undefined MAPKs.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AE003732; -; NOT ANNOTATED_CDS.
CC FlyBase; FBgn0046689; Tak11.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002290; Ser_Thr_kinase.
CC Pfam; PF00069; pkinase; 1.

```

ProDom; PD000001; Prot kinase; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;  
 KW Tyrosine-protein kinase; ATP-binding.  
 FT DOMAIN 11 266 PROTEIN KINASE.  
 FT NP BIND 17 25 ATP (By similarity).  
 FT BINDING 38 38 ATP (By similarity).  
 FT ACT SITE 133 133 BY SIMILARITY.  
 SQ SEQUENCE 393 AA; 45238 MW; 43370C98490F0384 CRC64;

Alignment Scores:  
 Pred. No.: 3.03e-35 Length: 393  
 Score: 435.50 Matches: 91  
 Percent Similarity: 60.78% Conservative: 50  
 Best Local Similarity: 39.22% Mismatches: 80  
 Query Match: 34.78% Indels: 11  
 DB: 1 Gaps: 6

US-09-830-144-1\_COPY\_408\_1091 (1-684) x M3K7\_DROME (1-393)

QY 4 GAGCTTGGGAGTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGAGCGCTGC 63  
 Db 54 GluileThrHisLeuSerGluLeuAspHisGluValIleArgValIleGlyArgAla 73  
 QY 64 TTGAAT-----CCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTTTATATAAT 117  
 Db 74 SerAsnGlyLysLysAspTyrLeuLeuMetGluTyrLeuGluGlySerLeuHisAsn 93  
 QY 118 GTGCTCGATGGTCTGCTGAACCATTTATATATATATATATATATATATATATATAT 177  
 Db 94 TyrLeuTyrGlyAspAspLysTyrGlu---TyrThrValGluGlnAlaValArgTyrAla 112  
 QY 178 TTACAGTGTCTCCAGGAGTGGTGTATCTTCACAGATGCAACCCAAAGCGCTTAATCAC 237  
 Db 113 LeuGlnCysAlaLysAlaLeuAlaTyrLeuHisSerLeuAsp---ArgProIleValHis 131  
 QY 238 AGGAGCTCAAAACCAACCAACTACTGCTGTTGTCAGGGGGGACAGTCTTAATAATTTGT 297  
 Db 132 ArgAspIleLysProGlnAsnMetLeuLeuTyrAsnGlnHisGluAspLeuLysIleCys 151  
 QY 298 GATTTGGTACGCTGTGATTCACACATTCAGACATCAACCAATAACAGGGAGTGTCTGT 357  
 Db 152 AspPheGlyLeuAlaThrAspMetSerAsnAsnLysThrAspMetGlnGlyThrLeuArg 171  
 QY 358 TGGATGGACCTCAAGCTTTTGAAGTAGTAATACAGTGAAATGTCAGCTTCACG 417  
 Db 172 TyrMetAlaProGluAlaIleLysHisLeuLysTyrThrAlaLysCysAspValTyrSer 191  
 QY 418 TGGGTATTATTCTTTGGGAAGTGATAACGGTCGGAAACCTTTGATGAGATTGGTGGC 477  
 Db 192 PheGlyIleMetLeuTyrGluLeuMetThrArgGlnLeuProTyrSerHisLeuGluAsn 211  
 QY 478 CCA-----GCTTCCGAATCATGTGGCTGTTCATAATGGTACTCGACCACTG--- 528  
 Db 212 ProAsnSerGlnTyrAlaIleMetLysAlaIleSerSerGlyLysLeuProMetGlu 231  
 QY 529 -----ATAAAAATTACTAAGCCCATGAGAGCTGATGACGTGTGGTGGCTAAA 582  
 Db 232 AlaValArgSerAspCysProGluGlyIleLysGlnLeuMetGluCysCysMetAspIle 251  
 QY 583 GATCTCTCCAGCGGCTTCATGAGAGAAATTTGAAATATGACTACTGCTGATGCGG 642  
 Db 252 AsnProGluLysArgProSerMetLysGluIleGluLysPheLeuGlyGlu-----Gln 269  
 QY 643 TACTTTCCAGGACGATGAGCCATTACAGTATCTCT 678  
 Db 270 TyrGluSerGlyThrAspGluAspPheIleLysPro 281

RESULT 4

M3KA\_HUMAN

ID M3KA\_HUMAN

STANDARD; PRT; 954 AA.

AC Q02779; Q12761; Q14871;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 20-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Mitogen-activated protein kinase kinase 10 (EC 2.7.1.37)  
 DE (Mixed lineage kinase 2) (Protein kinase MST).  
 GN MAP3K10 OR MLK2 OR MST.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=96128179; PubMed=8536694;  
 RA Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,  
 RA Sutherland G.R., Simpson R.J.;  
 RT "Complete nucleotide sequence, expression, and chromosomal  
 RT localisation of human mixed-lineage kinase 2.";  
 RL Eur. J. Biochem. 234:492-500(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=95249256; PubMed=7731697;  
 RA Katoh M., Hirai M., Sugimura T., Terada M.;  
 RT "Cloning and characterization of MSR, a novel (putative)  
 RT serine/threonine kinase with SH3 domain.";  
 RL Oncogene 10:1447-1451(1995).  
 RN [3]  
 RP SEQUENCE OF 244-480 FROM N.A.  
 RC TISSUE=Colon epithelium;  
 RX MEDLINE=9338756; PubMed=8477742;  
 RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;  
 RT "Identification of a new family of human epithelial protein kinases  
 RT containing two leucine/isoleucine-zipper domains.";  
 RL Eur. J. Biochem. 213:701-710(1993).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE SUBFAMILY.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X90846; CAA62351.1; -  
 DR EMBL; Z48615; CAA88531.1; -  
 DR PIR; S68178; S68178.  
 DR HSP; P11362; 1FGK.  
 DR Genew; HGNC:6849; MAP3K10.  
 DR MIM; 600137; -  
 DR GO; GO:0006917; P:induction of apoptosis; TAS.  
 DR GO; GO:0007254; P:JNK cascade; TAS.  
 DR GO; GO:0007165; P:signal transduction; TAS.  
 DR InterPro; IPR000719; Ser\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

DR PROSITE; PS50002; SH3; 1.  
 KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;  
 FT ATP-binding; SH3 domain.  
 FT DOMAIN 2 5 POLY-GLU.  
 FT DOMAIN 16 81 SH3.  
 FT DOMAIN 98 360 PROTEIN KINASE.  
 FT NP\_BIND 104 112 ATP (BY SIMILARITY).  
 FT BINDING 125 125 ATP (BY SIMILARITY).  
 FT ACT\_SITE 222 222 BY SIMILARITY.  
 FT DOMAIN 384 405 LEUCINE-ZIPPER 1 (BY SIMILARITY).  
 FT DOMAIN 419 440 LEUCINE-ZIPPER 2 (BY SIMILARITY).  
 FT DOMAIN 449 463 ARG/LYS-RICH (BASIC).  
 FT CONFLICT 462 464 SRL -> AV (IN REF. 2).  
 FT CONFLICT 465 480 LKLRGGSHLSISGF -> AQAGRQHPQPALWL (IN REF. 3).  
 FT CONFLICT 471 471 G -> S (IN REF. 2).  
 FT CONFLICT 807 807 G -> R (IN REF. 2).  
 FT CONFLICT 818 818 V -> A (IN REF. 2).  
 SQ SEQUENCE 954 AA; 103623 MW; 538F4AAA559B0ABA CRC64;  
 Alignment Scores:  
 Pred. No.: 6,44e-31 Length: 954  
 Score: 392.50 Matches: 88  
 Percent Similarity: 56.68% Conservatives: 35  
 Best Local Similarity: 40.55% Mismatches: 77  
 Query Match: 31.35% Indels: 17  
 DB: 1 Gaps: 6

US-09-830-144-1\_COPY\_408\_1091 (1-684) x M3KA\_HUMAN (1-954)

QY 4 GAGCTTCGGCAGTATATCCCGTGTGCAACCATCTTAATTTGTAAGCTTTATGGAGCTGC 63  
 Db 145 GLuAlaArgLeuPheGlyAlaLeuGluHisProAsnIleLeuAlaLeuArgGlyAlaCys 164  
 QY 64 TTGAATCCA-----GTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAAT 117  
 Db 165 LeuAsnProHisLeuCysLeuValMetGluTyrAlaArgGlyGlyAlaLeuSerArg 184  
 QY 118 GTGCTGATGCTGTGAACCATGTCATATATATCTGCTGCCACCAAGTGGTGTGT 177  
 Db 185 ValLeuAlaGlyArgValProHisVal-----LeuValAsnTyrAla 200  
 QY 178 TTACAGTGTTCACAGGAGTGGCTATCTTACAGCATGCAACCAAGCGCTAATTCAC 237  
 Db 201 ValGlnValAlaArgGlyMetAsnTyrLeuHisAsnAspAlaProValProIleHis 220  
 QY 238 AGGACCTGAAACCAACCAACTTACTGCTGTT-----GCAGGG 276  
 Db 221 ArgAspLeuLysSerIleAsnIleLeuGluAlaIleGluAsnHisAsnLeuAla 240  
 QY 277 GGGACAGTTCTAAATTTGATTTGGTACACCTGTGAC---ATTCAGACACACATG 333  
 Db 241 AspThrValLeuLysIleThrAspPheGlyLeuAlaArgGluTyrHisLysThrLys 260  
 QY 334 ACCAATACAGGGAGTGTGCTGTGATGTCACCTGACCTGAGCTTTTGAAGTAGTAATAC 393  
 Db 261 MetSerAlaAlaGlyThrTyrAlaTyrMetAlaProGluValIleArgLeuSerLeuPhe 280  
 QY 394 AGTGAAAAATGTGACGCTTTCAGCTGGGATATTCTTTGGGAAGTGAATACGCGTCGG 453  
 Db 281 SerLysSerSerAspValTyrSerPheGlyValLeuLeuTyrGluLeuLeuThrGlyGlu 300  
 QY 454 AAACCTTTTGATGATGTTGGGCCAGCTTTCGGAATCATGTGGCTGTT---CATAA 510  
 Db 301 ValProTyrArgGluIle-----AspAlaLeuAlaValAlaTyrGlyValAlaMetAsn 318  
 QY 511 GGTACTGCACCACTGATATAAAATTTACTAAGCCCATGAGAGCTGATGACTCGT 570  
 Db 319 LysLeuThrLeuProIleProSerThrCysProGluProPheAlaArgLeuGluGlu 338  
 QY 571 TGTGGTCTAAGATCTCTCCAGCGCTTCAATGGAGGAAATGTGAAA 621  
 Db 339 CysTrpAspProAspProHisGlyArgProAspPheGlySerIleLeuLys 355

RESULT 5  
 M3K9\_HUMAN STANDARD; PRT; 394 AA.  
 ID M3K9\_HUMAN  
 AC P80132;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Mitogen-activated protein kinase kinase 9 (EC 2.7.1.-) (Mixed  
 DE lineage kinase 1) (Fragment).  
 GN MAP3K9 OR MLK1 OR PRKE1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon epithelium;  
 RX MEDLINE=93238756; PubMed=8477742;  
 RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;  
 RT "Identification of a new family of human epithelial protein kinases  
 RT containing two leucine/isoleucine-zipper domains.";  
 RL Eur. J. Biochem. 213:701-710(1993).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF  
 CC COLONIC, BREAST AND OESOPHAGEAL ORIGIN.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 DR PIR; S32467; J00229.  
 DR HSSP; P12931; IPMK.  
 DR Genew; HGNC:6861; MAP3K9.  
 DR MIM; 600136; -  
 DR GO; GO:0005524; F:ATP binding activity; NAS.  
 DR GO; GO:0004708; F:MAP kinase activity; NAS.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; NAS.  
 DR InterPro; IPR00719; Prot kinase.  
 DR InterPro; IPR002290; Ser Thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE\_DOM; 1.  
 KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;  
 KW ATP-binding.  
 FT NON\_TER 1 1  
 FT DOMAIN 3 271 PROTEIN KINASE.  
 FT NP\_BIND 9 17 ATP (BY SIMILARITY).  
 FT BINDING 30 30 ATP (BY SIMILARITY).  
 FT ACT\_SITE 127 127 BY SIMILARITY.  
 FT DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).  
 FT DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).  
 FT DOMAIN 354 368 ARG/LYS-RICH (BASIC).  
 SQ SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;  
 Alignment Scores:  
 Pred. No.: 4.26e-28 Length: 394  
 Score: 363.50 Matches: 83  
 Percent Similarity: 54.09% Conservatives: 36  
 Best Local Similarity: 37.73% Mismatches: 84  
 Query Match: 29.03% Indels: 17  
 DB: 1 Gaps: 6

US-09-830-144-1\_COPY\_408\_1091 (1-684) x M3K9\_HUMAN (1-394)

QY 4 GAGCTTCGGCAGTATATCCCGTGTGCAACCATCTTAATTTGTAAGCTTTATGGAGCTGC 63  
 Db 50 GLuAlaArgLeuPheAlaMetLeuLysHisProAsnIleLeuAlaLeuArgGlyValCys 69  
 QY 64 TTGAATCCA-----GTGTGCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAAT 117  
 Db 70 LeuLysGluProAsnLeuCysLeuValMetGluPheAlaArgGlyGlyProLeuAsnArg 89

118 GTCTCATGGTGTGACCAATTCGATATTAATCTGCTGCCACGCAATGAGTGTGT 177  
 90 ValLeuSerGlyLysArgLysProAspLeu-----LeuValAsnTrpAla 105  
 178 TTACAGTGTTCACAGGAGTGGTATCTTCACAGCATGCAACCAAGCGCTAATTCAC 237  
 106 ValGlnIleAlaArgGlyMetAsnTrpLeuHisAspGluAlaIleValProIleHis 125  
 238 AGGACCTGAACACCAACCAATTCAGTGTGT-----GCAGGG 276  
 126 ArgAspLeuLysSerAsnIleLeuIleGlnLysValGluAsnGlyAspLeuSer 145  
 277 GGGACAGTCTATAAATTTGTGATTTGTGTACAGCTGTGC---ATTACACACACATG 333  
 146 AsnLysIleLeuLysIleThrAspPheGlyLeuAlaArgLysTrpHisArgThrLys 165  
 334 ACCAATAACAGGGAGTGTCTGTCAGGACCTGAGTGTTCAGGTAGTATTCAC 393  
 166 MetSerAlaAlaGlyThrTrpAlaTrpMetAlaProGluValIleArgAlaSerMetPhe 185  
 394 AGTGAATAATGTGACGTCTTCAGCTGGGTGATTTATCTTTGGAGTGTAAACGCTGG 453  
 186 SerLysGlySerAspValTrpSerTrpGlyValLeuLeuTrpGlnLeuLeuThrGly 205  
 454 AAACCCCTTTGATGAGATTTGTGCTGCCAGCTTTCCGAATCATGTGGGTGT---CATAAT 510  
 206 ValProPheArgGlyIleAspGly-----LeuArgValAlaTrpGlyValAlaMetAsn 223  
 511 GGTACTCGACCACTGATATAAATTTACCTAAGCCATTCAGAGCTGTGATCGT 570  
 224 LysLeuAlaLeuProIleProSerThrCysProGluProPheAlaLysLeuMetGluAsp 243  
 571 TGTGTCTTAAGATCTTCCAGCGCTTCATGAGGAAATTCGAGAAATGCAATGACT 630  
 244 CysTrpAsnProAspProHisArgProSerPheThrAsnIleLeuAspGlnLeuThr 263

RESULT 6  
 M3KC RAT  
 ID M3KC RAT STANDARD; PRT; 888 AA.  
 AC Q63796;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mitogen-activated protein kinase kinase kinase 12 (BC 2.7.1.37)  
 DE (MAPK-upstream kinase) (MUK).  
 GN MAP3K12 OR MUK.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96226099; PubMed=8637721;  
 RA Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.;  
 RT "Activation of the JNK pathway by distantly related protein kinases,  
 RT MEK and MUK."  
 RL Oncogene 12:641-650(1996).  
 CC -!- FUNCTION: May be an activator of the JNK/SAPK pathway.  
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in  
 CC vitro.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- COFACTOR: Magnesium.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By  
 CC similarity).  
 CC -!- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol  
 CC under basal conditions and dephosphorylated when membrane-  
 CC associated (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MAP KINASE KINASE KINASE SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL; D49785; BAA08621.1; -  
 DR HSSP; P12931; LPK.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_Thr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR Transferrase; Serine/threonine-protein kinase; ATP-binding;  
 KW Phosphorylation; Magnesium; Membrane.  
 FT DOMAIN 158 399 PROTEIN\_KINASE.  
 FT NP\_BIND 164 172 ATP (BY SIMILARITY).  
 FT BINDING 185 185 ATP (BY SIMILARITY).  
 FT ACT\_SITE 269 269 BY SIMILARITY.  
 FT DOMAIN 56 62 POLY-GLY.  
 FT DOMAIN 668 671 POLY-PRO.  
 FT DOMAIN 698 701 POLY-PRO.  
 FT DOMAIN 753 758 POLY-GLU.  
 FT SEQUENCE 888 AA; 96307 MW; 52AD964006BAE149 CRC64;  
 -----  
 Alignment Scores:  
 Pred. No.: 3e-27 Length: 888  
 Score: 355.50 Matches: 75  
 Percent Similarity: 55.09% Conservative: 44  
 Best Local Similarity: 34.72% Mismatches: 80  
 Query Match: 28.39% Indels: 17  
 DB: 1 Gaps: 8  
 -----  
 US-09-830-144-1\_COPY\_408\_1091 (1-684) x M3KC\_RAT (1-888)  
 QY 4 GAGCTTCGCGAGTATCCCGTGTGAACATCTCTAATATGTAAAGCTTTATGAGCGCTGC 63  
 Db 194 AspileLysHisLeuArgLysHisProAsnIleIleThrPhelGlyValCys 213  
 QY 64 TTGAATCCA-----GTGTGTCTTGTGATGGAATATGCTGAAGGGGCTTTATATAAT 117  
 Db 214 ThrGlnAlaProCysTyrCysIleLeuMetGluPheCysAlaGlnGlyLysLeuThrGlu 233  
 QY 118 GTGCTGATGTGCTGAACCATTTGCCATATTAATCTGCTGCCACCAATGAGTGTGT 177  
 Db 234 ValLeuArgAlaGlyArgProVal-----ThrProSerLeuLeuValAspTrpSer 250  
 QY 178 TTACAGTGTTCACAGGAGTGTCTTATCTTCACAGCATGCAACCAAGCGCTAATTCAC 237  
 Db 251 MetGlyIleAlaGlyLysMetAsnTrpLeuHisLys-----IleIleHis 267  
 QY 238 AGGACCTGAACACCAACCAATTCAGTGTGTGAGGGGGGACAGTCTTCAAAATTCGT 297  
 Db 268 ArgAspLeuLysSerProAsn---MetLeuIleThrTyrAspValValLysIleSer 286  
 QY 298 GATTTGTGTACAGCTGTGACATTCAGACACACATGACCAATACAAAG-----GGGAGT 351  
 Db 287 AspPheGlyThrSerLysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThr 306  
 QY 352 GCTGTCTGATGGACCTGAGTTTGAAGTAGTAAATACAGTGAATAATGTCACGTC 411  
 Db 307 ValAlaTrpMetAlaProGluValIleArgAsnGluProValSerGluLysValAspIle 326  
 QY 412 TTCACCTGGGTATTTATCTTTGGAGTGAATAACGCTCGGAAACCCCTTTCATGAGATT 471  
 Db 327 TrpSerPheGlyValValLeuTrpGluLeuLeuThrGlyGluIleProTyrLysAspVal 346  
 QY 472 GGTGCCCGCAGTTCCTCGAATCATGTTGGCTGT---CATAATGCTACTGACACCACTG 528  
 Db 472 GGTGCCCGCAGTTCCTCGAATCATGTTGGCTGT---CATAATGCTACTGACACCACTG 528



```

Db 347 AspSerSerAla-----IleIleTTPGlyValGlySerAsnSerLeuHisLeuProVal 364
QY 529 ATAAAAAATTACTAAGCCCATGAGCGCTGATGACTCGTTGGTGTCTAAGATCCT 588
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 365 ProSerSerCysProAspGlyPheLysIleLeuLeuArgGlnCysTrpAsnArgLysPro 384
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 589 TCCGAGCGCCCTTCAATGAGGAGAAATGTGAAATATGACTCCTTG 636
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 385 ArgAsnArgProSerPheArgGln-----IleLeuLeuHisLeu 397

RESULT 7
M3KC_HUMAN
ID M3KC_HUMAN STANDARD; PRT; 859 AA.
AC Q12852;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
DE (leucine-zipper protein kinase) (ZPK).
GN MAP3K12 OR ZPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=teratocarcinoma;
RX MEDLINE=94311945; PubMed=8037767;
RA Reddy U.R.; Pleasure D.;
RT "Cloning of a novel putative protein kinase having a leucine zipper
RL domain from human brain.";
RL Biochem. Biophys. Res. Commun. 202:613-620(1994).
CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- COFACTOR: Magnesium.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.
CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U07358; AAA67343.1; -.
CC HSPSP; P12931; 1PMK.
CC Genew; HGNC:6851; MAP3K12.
CC MTM; 600447; -.
CC DR GO; GO:0005737; C:cytoplasm; TAS.
CC DR GO; GO:0005886; C:plasma membrane; TAS.
CC DR GO; GO:0007254; P:JNK cascade; TAS.
CC DR InterPro; IPR000719; Prot_kinase.
CC DR InterPro; IPR002290; Ser_thr_kinase.
CC DR InterPro; IPR001245; Tyr_kinase.
CC DR Pfam; PF00069; pkinase; 1.
CC DR PRINTS; PR00109; TYRKINASE.
CC DR ProDom; PD000001; Prot_kinase; 1.
CC DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
CC DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation; Magnesium; Membrane.
CC DOMAIN 125 366 PROTEIN KINASE.

```

```

FT NP_BIND 131 139 ATP (BY SIMILARITY).
FT BINDING 152 152 ATP (BY SIMILARITY).
FT ACT_SITE 236 236 BY SIMILARITY.
FT DOMAIN 665 668 POLY-PRO.
FT DOMAIN 720 725 POLY-GLU.
SQ SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;

Alignment Scores:
Pred. No.: 4.72e-27 Length: 859
Score: 353.50 Matches: 75
Percent Similarity: 54.63% Conservative: 43
Best Local Similarity: 34.72% Mismatches: 81
Query Match: 28.23% Indels: 17
DB: Gaps: 8

US-09-830-144-1_COPY_408_1091 (1-684) x M3KC_HUMAN (1-859)
QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTTATGAGCCTGC 63
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 161 AspileyHisLeuArgLysLeuLysHisProAsnIleIleThrPheLysGlyValCys 180
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 64 TTGAATCCA-----GTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTATATAT 117
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 181 ThrGlnAlaProCysTyrCysIleLeuMetGluPheCysAlaGlnGlyGlnLeuTyrGlu 200
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 118 GTGCTGCATGGTCTGAACCATTCATATATATCTGCTGCCACCGCAATGAGTGTGT 177
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 201 ValLeuArgAlaGlyArgProVal-----ThrProSerLeuLeuValAspTrpSer 217
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 178 TTACAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGATGCAACCCAAAGCGCTAATTCAC 237
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 218 MetGlyIleAlaGlyGlyMetAsnTyrLeuHisLeuHisLys-----IleIleHis 234
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 238 AGGACCTGAAACACCAACAACTTACTGCTGGTTCAGAGGGGGGACAGTCTTAAATTTGT 297
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 235 ArgAspLeuLysSerProAsn---MetLeuIleThrTyrAspAspValValLysIleSer 253
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 298 GATTTTGTACAGCTGTGACATTCAGACACACATGACCAATACAAAG-----GGGAGT 351
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 254 AspPheGlyThrSerLysGluLeuSerAspLysSerThrLysMetSerPheAlaGlyThr 273
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 352 GCTGCTTGGATGCGCTGAAAGTTTTGAAGGTAGTAGTAATTACAGTGAATAATGTGAGTGC 411
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 274 ValAlaTrpMetAlaProGluValIleArgAsnGluProValSerGluLysValAspIle 293
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 412 TTCAGCTGGGGTATATTCTTTCGGAAGTAGTAAACCGCTCGGAAACCTTTGATGAGATT 471
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 294 TrpSerPheGlyValValLeuTrpGluLeuLeuThrGlyGluIleProTyrLysAspVal 313
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 472 GTGGGCCAGCTTCCGAATCATGTGGGCTGTT---CATTAATGGTACTCGACCACTG 528
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 314 AspSerSerAla-----IleIleTrpGlyValGlySerAsnSerLeuHisLeuProVal 331
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 529 ATAAAAAATTACCTAAGCCCATGAGCTGATGACTCGTTGTTGGTCTAAAGATCCT 588
    :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 332 ProSerSerCysProAspGlyPheLysIleLeuLeuArgGlnCysTrpAsnSerLysPro 351
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 589 TCCGAGCGCCCTTCAATGAGGAGAAATGTGAAATATGACTCCTTG 636
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 352 ArgAsnArgProSerPheArgGln-----IleLeuLeuHisLeu 364
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 8
M3KC_MOUSE
ID M3KC_MOUSE STANDARD; PRT; 888 AA.
AC Q60700; P70286;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
DE (leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing
DE kinase) (DLK).
GN MAP3K12 OR ZPK.
OS Mus musculus (Mouse).

```



DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).  
GN PYKB OR DPYK2.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
ON NCBI\_TaxID=44689;  
RX [1]  
RY SEQUENCE FROM N.A.  
RZ MEDLINE=90287147; PubMed=1972546;  
RA Tan J.L., Spudich J.A.;  
RT "Developmentally regulated protein-tyrosine kinase genes in  
RT Dictyostelium discoideum";  
RL Mol. Cell. Biol. 10:3578-3583(1990).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M33784; AAA33203.1; --  
CC PIR; B35670; B35670.  
CC HSP; P08631; 1AD5.  
CC Dictyob; DD03011; PykB.  
CC InterPro; IPR00719; Prot\_kinase.  
CC InterPro; IPR001245; Tyr\_kinase.  
CC Pfam; PF00069; pkinase; 1.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
CC Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.  
KW NON\_TPR 1 1  
FT DOMAIN 108 381 PROTEIN\_KINASE  
FT NP\_BIND 114 122 ATP (BY SIMILARITY).  
FT BINDING 135 135 ATP (BY SIMILARITY).  
FT ACT\_SITE 232 232 BY SIMILARITY.  
SQ SEQUENCE 410 AA; 46386 MW; E93918B605B9AEC1 CRC64;

Alignment Scores:  
Pred. No.: 7,34e-26 Length: 410  
Score: 341.00 Matches: 74  
Percent Similarity: 54.21% Conservative: 42  
Best Local Similarity: 34.58% Mismatches: 84  
Query Match: 27.24% Indels: 14  
DB: 1 Gaps: 7

US-09-830-144-1\_COPY\_408\_1091 (1-684) x KYK2\_DICDI (1-410)

QY 4 GAGCTCGGAGTTATCCCTGTAACCATCTTAATATTGTAAGCTTTATGGAGCCTGC 63  
Db 153 GluValGlnAsnLeuLysGlyAsnHisGlnAsnIleValMetPheIleGlyAlaCys 172  
QY 64 TTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGCTG 123  
Db 173 TyrLysProAlaCysIleIleThrGluTy-MetAlaGlyGlySerLeuTyrAsnIleLeu 192  
QY 124 CAT-----GGTGTGACCATTCGCATATTATCTACTGCTGCCAGCAATGAGTGG 174  
Db 193 HisAsnProAsnSerSerThrProLysValLysTyrSerPheProLeuValLysMet 212  
QY 175 TGTTTACAGTCTTCCCAAGGAGTGGCTTATCTTACAGCATGCAACCAAGCGCTAATT 234  
Db 213 AlaThrAspMetAlaLeuGlyLeuLeuHisHisSerIle-----ThrIleVal 229  
QY 235 ACAGGAGCCTGAACACCAACTTACTGCTGCTGAGGGGGAGCAGTTCTAAAAATT 294  
Db 230 HisArgAspLeuThrSerGlnAsnIleLeuLeuAspGluLeuGlyAsnIle---LysIle 248

QY 295 TGTGATTTTGCTACAGCCTGTGAC-----ATTGAGACACACATGACCAATAACAAGGGG 348  
Db 249 SerAspPheGlyLeuSerAlaGluLysSerArgGluGlySerMetThrMetThrAsnGly 268  
QY 349 -----AGTGCTGCTTGGATGGCACCTGCAAGTCTTTTGAAGGT---AGTAATTACAGT 396  
Db 269 GlyIleCysAsnProAsnArgTyrArgProGluLeuLeuThrLysAsnLeuGlyHisTyrSer 288  
QY 397 GAAAAATGTGACGCTTCTCAGCTGGGTATTATCTTTGGGAAGTGAATACCGCTCGGAAA 456  
Db 289 GluLysValAspValTyrCysPheSerLeuValTyrGluIleLeuThrGlyGluIle 308  
QY 457 CCCTTTGATGAGATTGGTGGCCCGACGCTTCGGAATCATGTGGCTGTTCATAATGGTACT 516  
Db 309 ProPheSerAspLeuAspGly---SerGlnArgSerAlaGlnValAlaTyrAlaGlyLeu 327  
QY 517 CGACCCACCTGATAAAATAATTTACCTTAAGCCCATGAGAGCCCTGATGACTCGTTGTGG 576  
Db 328 ArgProIleProGluTyrCysAspProGluLeuLysLeuLeuLeuThrGlnCysTyr 347  
QY 577 TCTAAGATCTTCCCGACGCCCTTCAATGAGGAGAAATTGTG 618  
Db 348 GluAlaAspProAsnAspArgProProPheThrTyrIleVal 361

RESULT 10  
CTRL\_ARATH  
ID\_CTRL\_ARATH STANDARD; PRT; 821 AA.  
AC Q05609;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Serine/threonine-protein kinase CTRL (EC 2.7.1.37).  
GN CTRL OR AT5G03730 OR F17C15150.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Eumatiophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia; TISSUE=Seedling;  
RC MEDLINE=93161417; PubMed=8431946;  
RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;  
RT "CTRL, a negative regulator of the ethylene response pathway in  
RT Arabidopsis, encodes a member of the raf family of protein kinases.";  
RL Cell 72:427-441(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RC MEDLINE=21016721; PubMed=11130714;  
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,  
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,  
RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,  
RA Nakazaki N., Naruo K., Okumura S., Shino S., Takeuchi C., Wada T.,  
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,  
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,  
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,  
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,  
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,  
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,  
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,  
RA Wagner-McPherson C., Wollam A., Yokum W., Bell M., Dedhia N.,  
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,  
RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,  
RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,  
RA Volckaert G., Mambutt R., Duesterhoeft A., Stiekema W., Pohl T.,  
RA Etian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,  
RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,  
RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,  
RA van Staveren M., Dirks W., Mooijman P., Klein Lankhorst R.,  
RA Weitzenecker T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,  
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,  
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,

RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;  
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis  
 RL thaliana";  
 CC Nature 408:823-826 (2000).  
 CC -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE  
 CC PATHWAY.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.  
 CC -1- MISCELLANEOUS: CTRL MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,  
 CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED  
 CC COTYLEDON GROWTH IS IMPAIRED.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC MIL/RAF SUBFAMILY.  
 CC  
 CC This SW-SS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; L08789; AAA32779.1; -;  
 CC EMBL; L08790; AAA32780.1; -;  
 CC EMBL; AL162506; CAB82938.1; -;  
 CC FIR; T48400; T48400.  
 CC InterPro; IPR000719; Prot kinase.  
 CC InterPro; IPR002290; Ser\_Thr\_kinase.  
 CC InterPro; IPR01245; Tyr\_kinase.  
 CC Pfam; PF00669; pkinase; 1.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC ProDom; PD000001; Prot kinase; 1.  
 CC SMART; SM00220; S\_TKc; 1.  
 CC PROSITE; PS00107; "PROTEIN KINASE ATP; 1."  
 CC PROSITE; PS00108; "PROTEIN KINASE ST; 1."  
 CC PROSITE; PS00111; "PROTEIN KINASE DOM; 1."  
 CC Transferrase; Serine/threonine-protein kinase; ATP-binding.  
 CC  
 CC FT DOMAIN 65 69 POLY-GLY.  
 CC FT DOMAIN 135 141 POLY-GLY.  
 CC FT DOMAIN 551 809 PROTEIN KINASE.  
 CC FT NP\_BIND 557 565 ATP (BY SIMILARITY).  
 CC FT BINDING 578 578 ATP (BY SIMILARITY).  
 CC FT ACT\_SITE 676 676 BY SIMILARITY.  
 CC FT MUTAGEN 596 596 E->K: IN CTRL-4; EXHIBITS ETHYLENE-  
 CC TREATED PHENOTYPE.  
 CC FT MUTAGEN 694 694 D->E: IN CTRL-1; EXHIBITS ETHYLENE-  
 CC TREATED PHENOTYPE.  
 CC SQ SEQUENCE 821 AA; 90306 MW; 2922D3DCD0CC15BC CRC64;  
 Alignment Scores:  
 Pred. No.: 5e-24 Length: 821  
 Score: 323.00 Matches: 74  
 Percent Similarity: 54.59% Conservative: 45  
 Best Local Similarity: 33.94% Mismatches: 91  
 Query Match: 25.80% Indels: 8  
 DB: 1 Gaps: 6  
 US-09-830-144-1\_COPY\_408\_1091 (1-684) x CTRL\_4PATH (1-821)  
 QY 4 GAGCTTCGGCAGTTATCCCGTGTGAACCACTCTTAATATTGTAAGCTTTATGAGCGCTGC 63  
 Db 596 GluValAlaLeuMetIysArgLeuArgHisProAsnIleValLeuPheMetGlyAlaVal 615  
 QY 64 TTGAATCCCA-----GTGGTCTGTGGATGATATGCTGAAGGGGCTCTTATATAT 117  
 Db 616 ThrGlnProProAsnLeuSerIleValThrGluTyrLeuSerArgGlySerLeuTyrArg 635  
 QY 118 GTGCTGCATGCTGCTGAACCACTTGCATATATTAATCTGCTGCCACCAATGAGTTGGTGT 177  
 Db 636 LeuLeuHisLysSerGlyAlaArgGluGlnLeuAspGluArgAlaArgSerMetAla 655  
 QY 178 TTACAGTGTCCAGGAGTGGCTTATCTTCACAGATGCAACCCAAACGGCTAATTCAC 237

Db 656 TyrAspValAlaLysGlyMetAsnTyrLeuHisAsnArgAsnPro----ProIleValHis 674  
 QY 238 AGGAGCTGAACACCAACCAACTTACTGCTGTTGCAGGGGGGACAGTTCTAAAAATTGT 297  
 Db 675 ArgAspLeuLysSerProAsnLeuValAspLysLysTyrThrVal---LysValCys 693  
 QY 298 GATTTTGGT---ACAGCTGTGACATTCAGACACACATGACCAATACAAAG-----GGG 348  
 Db 694 AspPheGlyLeuSerArgLeuLysAlaSerThrPheLeuSerSerLysSerAlaAlaGly 713  
 QY 349 AGTGCTGTGGATGACACTGAAGTTTGTGAAGGTAGTAATTACAGTGAATAATGTGAC 408  
 Db 714 ThrProGluTrpMetAlaProGluValLeuArgAspGluProSerAsnGluLysSerAsp 733  
 QY 409 GCTTTCAGCTGGGTATTATCTTTCGGAAGTCATAACGCTCGGAACCCCTTGTGATGAG 468  
 Db 734 ValTyrSerPheGlyValIleLeuTrpGluLeuAlaThrLeuGlnProTrpGlyAsn 753  
 QY 469 ATTTGGTGGCCAGCTTTCCGAATCATGTGGCTGTTCATAATGTTACTCGACCCACTG 528  
 Db 754 Leu---AsnProAlaGlnValValAlaAlaValGlyPheLysCysLysArgLeuGluIle 772  
 QY 529 ATAAATAATTTACTTAAGCCCTATGAGCCCTGATGACTGCTGTGGTGTAAAGATCCT 588  
 Db 773 ProArgAsnLeuAsnProGlnValAlaAlaIleIleGluGlyCysTrpThrAsnGluPro 792  
 QY 589 TCCACGCGCCTCAATGAGGAAATTTGAAATAATGACTCCTGATGTCGG 642  
 Db 793 TrpLysArgProSerPheAlaThrIleMetAspLeuLeuArgProIleLys 810  
 RESULT 11  
 RET MOUSE  
 ID RET MOUSE STANDARD; PRT; 1115 AA.  
 AC P35346;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Proto-oncogene tyrosine-protein kinase receptor ret precursor  
 DE (EC 2.7.1.112) (C-ret).  
 GN RET.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 RN (1)\_TaxID=10090;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93205390; PubMed=8455936;  
 RA Iwanoto T., Taniguchi M., Asai N., Ohkusu K., Nakashima I.,  
 RA Takahashi M.;  
 RT "cDNA cloning of mouse ret proto-oncogene and its sequence similarity  
 RT to the cadherin superfamily";  
 RL Oncogene 8:1087-1091 (1993).  
 [2]  
 RP INTERACTION WITH DOK2; DOK4 AND DOK5, PHOSPHORYLATION, AND MUTAGENESIS  
 OF TYR-1063.  
 RX MEDLINE=21363571; PubMed=11470823;  
 RA Grimm J., Sachs M., Britsch S., Di Cesare S., Schwarz-Romond T.,  
 RA Alitalo K., Birchmeier W.;  
 RT "Novel p62dok family members, dok-4 and dok-5, are substrates of the  
 RT c-ret receptor tyrosine kinase and mediate neuronal  
 RT differentiation";  
 RL J. Cell Biol. 154:345-354 (2001).  
 CC -1- FUNCTION: Probable receptor with tyrosine-protein kinase activity;  
 CC important for development.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBUNIT: Phosphorylated form interacts with the PBT domain of  
 CC DOK2, DOK4 and DOK5.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: Expressed in peripheral nerve cells and  
 CC hematopoietic cells.  
 CC -1- PTM: Phosphorylated.  
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.



kinase with two potential transmembrane domains.";  
 [3]  
 RP SEQUENCE OF 588-1063 FROM N.A.  
 RX MEDLINE=97257826; PubMed=3037315;  
 RA Takahashi M., Cooper G.M.;  
 RT "ret transforming gene encodes a fusion protein homologous to  
 RT tyrosine kinases.";  
 RL Hum. Mol. Biol. 7:1378-1385(1987).  
 [4]  
 RP TYROSINE AUTOPOHOSPHORYLATION, AND MUTAGENESIS OF TYR-1015 AND  
 RP TYR-1062.  
 RX MEDLINE=20513733; PubMed=11061555;  
 RA Salvatore D., Barone M.V., Salvatore G., Melillo R.M., Chiappetta G.,  
 RA Mineo A., Penzi G., Vecchio G., Fusco A., Santoro M.;  
 RT "Tyrosines 1015 and 1062 are in vivo autophosphorylation sites in ret  
 RT and ret-derived oncoproteins.";  
 RL J. Clin. Endocrinol. Metab. 85:3898-3907(2000).  
 [5]  
 RP REVIEW ON HSCR VARIANTS.  
 RP MEDLINE=98023959; PubMed=9359036;  
 RA Hofstra R.M.W., Osinga J., Buys C.H.C.M.;  
 RT "Mutations in Hirschsprung disease: when does a mutation contribute to  
 RT the phenotype.";  
 RL Eur. J. Hum. Genet. 5:180-185(1997).  
 [6]  
 RP REVIEW ON VARIANTS.  
 RP MEDLINE=97220587; PubMed=9067749;  
 RA Eng C., Mulligan L.M.;  
 RT "Mutations of the RET proto-oncogene in the multiple endocrine  
 RT neoplasia type 2 syndromes, related sporadic tumours, and  
 RT Hirschsprung disease.";  
 RL Hum. Mutat. 9:97-109(1997).  
 [7]  
 RP VARIANTS MEN2A/FMTC TRP-611; SER-618; ARG-620; TYR-620 AND ARG-634.  
 RX MEDLINE=93372843; PubMed=8103403;  
 RA Donis-Keller H., Dou S., Chi D., Carlson K.M., Toshima K.,  
 RA Laitmore T.C., Howe J.R., Moley J.F., Goodfellow P., Wells S.A. Jr.;  
 RT "Mutations in the RET proto-oncogene are associated with MEN 2A and  
 RT FMTC.";  
 RL Hum. Mol. Genet. 2:851-856(1993).  
 [8]  
 RP VARIANTS MEN2A GLY-618; 632-ASP-VAL-ARG-634; GLX-634; PHE-634; TYR-634  
 RP AND SER-634.  
 RX MEDLINE=93275414; PubMed=8099202;  
 RA Mulligan L.M., Kwok J.B.J., Healey C.S., Elsdon M.J., Eng C.,  
 RA Gardner E., Love D.R., Mole S.E., Moore J.K., Papi L., Ponder M.A.,  
 RA Telenius H., Tunncliffe A., Ponder B.A.J.;  
 RT "Germline mutations of the RET proto-oncogene in multiple endocrine  
 RT neoplasia type 2A.";  
 RL Nature 363:458-460(1993).  
 [9]  
 RP VARIANTS HSCR PRO-40; LEU-399; GLN-762; PRO-765; GLN-897; GLY-972 AND  
 RP LEU-973.  
 RX MEDLINE=95219414; PubMed=7704557;  
 RA Yin L., Barone V., Seri M., Bolino A., Bocciaardi R., Ceccherini I.,  
 RA Pasini B., Tocco T., Lerone M., Cywes S., Moore S.,  
 RA Vanderwinden J.-M., Abramowicz M.J., Kristofferson U., Larsson L.T.,  
 RA Hamel B.C.J., Silengo M., Martucciello G., Romeo G.;  
 RT "Heterogeneity and low detection rate of RET mutations in Hirschsprung  
 RT disease.";  
 RL Eur. J. Hum. Genet. 2:272-280(1994).  
 [10]  
 RP VARIANT MEN2B THR-918.  
 RX MEDLINE=94272459; PubMed=7911697;  
 RA Eng C., Smith D.P., Mulligan L.M., Nagai M.A., Healey C.S.,  
 RA Ponder M.A., Gardner E., Scheumann G.F., Jackson C.E., Tunncliffe A.,  
 RA Ponder B.A.J.;  
 RT "Point mutation within the tyrosine kinase domain of the RET  
 RT proto-oncogene in multiple endocrine neoplasia type 2B and related  
 RT sporadic tumours.";  
 RL Hum. Mol. Genet. 3:237-241(1994).  
 [11]

VARIANTS MEN2A/FMTC ARG-618; SER-618; PHE-620; ARG-620; PHE-634;  
 RP GLY-634 AND TYR-634.  
 RX MEDLINE=94348513; PubMed=79151565;  
 RA Xue F., Yu H., Maurer L.H., Memoli V.A., Nutile-Mcmenemy N.,  
 RA Schuster M.K., Browden D.W., Mao J.-I., Noll W.W.;  
 RT "Germline RET mutations in MEN 2A and FMTC and their detection by  
 RT simple DNA diagnostic tests.";  
 RL Hum. Mol. Genet. 3:635-638(1994).  
 [12]  
 RP VARIANTS FMTC/MEN2A TYR-609; ARG-618; SER-618 AND SER-620.  
 RX MEDLINE=95152521; PubMed=7849720;  
 RA Blaugrund J.E., Johns M.M. Jr., Eby Y.J., Ball D.W., Baylin S.B.,  
 RA Hruban R.H., Sidransky D.;  
 RT "RET proto-oncogene mutations in inherited and sporadic medullary  
 RT thyroid cancer.";  
 RL Hum. Mol. Genet. 3:1895-1897(1994).  
 [13]  
 RP VARIANTS FMTC, AND VARIANTS MEN2A.  
 RP MEDLINE=95179108; PubMed=7874109;  
 RA Schuffenecker I., Billaud M., Calender A., Chambe B., Ginot N.,  
 RA Calmettes C., Modigliani E., Lenoir G.M.;  
 RT "RET proto-oncogene mutations in French MEN 2A and FMTC families.";  
 RL Hum. Mol. Genet. 3:1939-1943(1994).  
 [14]  
 RP VARIANT HSCR TRP-609, VARIANT HSCR/MEN2A ARG-618, AND VARIANT  
 RP HSCR/FMTC ARG-620.  
 RX MEDLINE=95181155; PubMed=7881414;  
 RA Mulligan L.M., Eng C., Attie T., Lyonnet S., Marsh D.J., Hyland V.J.,  
 RA Robinson B.G., Frilling A., Verellen-Dumoulin C., Safar A.,  
 RA Venter D.J., Munnich A., Ponder B.A.J.;  
 RT "Diverse phenotypes associated with exon 10 mutations of the RET  
 RT proto-oncogene.";  
 RL Hum. Mol. Genet. 3:2163-2167(1994).  
 [15]  
 RP VARIANT MEN2B THR-918.  
 RX MEDLINE=94159102; PubMed=7906866;  
 RA Hofstra R.M.W., Landsvater R.M., Ceccherini I., Stulp R.P.,  
 RA Stelwagen T., Luo Y., Pasini B., Hoepfner J.W.M., van Amstel H.K.P.,  
 RA Romeo G., Lips C.J.M., Buys C.H.C.M.;  
 RT "A mutation in the RET proto-oncogene associated with multiple  
 RT endocrine neoplasia type 2B and sporadic medullary thyroid  
 RT carcinoma.";  
 RL Nature 367:375-376(1994).  
 [16]  
 RP VARIANTS HSCR PRO-765; GLN-897 AND GLY-972.  
 RX MEDLINE=94159103; PubMed=8114938;  
 RA Romeo G., Ronchetto P., Luo Y., Barone V., Seri M., Ceccherini I.,  
 RA Pasini B., Bocciaardi R., Lerone M., Kaarlainen H., Martucciello G.;  
 RT "Point mutations affecting the tyrosine kinase domain of the RET  
 RT proto-oncogene in Hirschsprung's disease.";  
 RL Nature 367:377-378(1994).  
 [17]  
 RP VARIANTS HSCR LEU-32; LEU-64; GLN-330 AND LEU-393.  
 RX MEDLINE=94159104; PubMed=8114939;  
 RA Edery P., Lyonnet S., Mulligan L.M., Pelet A., Dow E., Abel L.,  
 RA Holder S., Nihoul-Fkete C., Ponder B.A.J., Munnich A.;  
 RT "Mutations of the RET proto-oncogene in Hirschsprung's disease.";  
 RL Nature 367:378-380(1994).  
 [18]  
 RP VARIANT MEN2B THR-918.  
 RX MEDLINE=94151373; PubMed=7906417;  
 RA Carlson K.M., Dou S., Chi D., Scavarda N., Toshima K., Jackson C.E.,  
 RA Wells S.A. Jr., Goodfellow P.J., Donis-Keller H.;  
 RT "Single missense mutation in the tyrosine kinase catalytic domain of  
 RT the RET protooncogene is associated with multiple endocrine neoplasia  
 RT type 2B.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1579-1583(1994).  
 [19]  
 RP VARIANTS MTC; FMTC; MEN2A AND MEN2B.  
 RX MEDLINE=96223053; PubMed=8625130;  
 RA Kommeth P., Kunz E.K., Matias-Guiu X., Hiort O., Christiansen G.,  
 RA Colomer A., Roth J., Heitz P.U.;  
 RT "Analysis of RET protooncogene point mutations distinguishes heritable



CC IsoId=P24604-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P24604-2; Sequence=VSP\_005012, VSP\_005015;  
 CC Name=3; Synonyms=RecLib;  
 CC IsoId=P24604-3; Sequence=VSP\_005016;  
 CC Name=4; Synonyms=RecIIa;  
 CC IsoId=P24604-4; Sequence=VSP\_005015, VSP\_005016;  
 CC Name=5; Synonyms=RecII;  
 CC IsoId=P24604-5; Sequence=VSP\_005015;  
 CC Name=6; Synonyms=RecI;  
 CC IsoId=P24604-6; Sequence=VSP\_005013, VSP\_005014;  
 CC -!- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN LIVER. EXPRESSION  
 CC IS ALSO SEEN IN THE HEMATOPOIETIC CELLS SUCH AS BONE MARROW,  
 CC THYMUS AND SPLEEN. LOWER EXPRESSION SEEN IN THE HEART, KIDNEY AND  
 CC OVARY.  
 CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. TEC  
 CC SUBFAMILY.  
 CC -!- SIMILARITY: Contains 1 SH2 domain.  
 CC -!- SIMILARITY: Contains 1 SH3 domain.  
 CC -!- SIMILARITY: Contains 1 PH domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; S53716; AAA13515.2; -;  
 CC EMBL; AF071938; AAD43404.1; -;  
 CC EMBL; AF071936; AAD43404.1; JOINED.  
 CC EMBL; AF071937; AAD43404.1; JOINED.  
 CC EMBL; AF071946; AAD43402.1; -;  
 CC EMBL; AF071936; AAD43402.1; JOINED.  
 CC EMBL; AF071937; AAD43402.1; JOINED.  
 CC EMBL; AF071938; AAD43402.1; JOINED.  
 CC EMBL; AF071939; AAD43402.1; JOINED.  
 CC EMBL; AF071940; AAD43402.1; JOINED.  
 CC EMBL; AF071941; AAD43402.1; JOINED.  
 CC EMBL; AF071942; AAD43402.1; JOINED.  
 CC EMBL; AF071943; AAD43402.1; JOINED.  
 CC EMBL; AF071944; AAD43402.1; JOINED.  
 CC EMBL; AF071945; AAD43402.1; JOINED.  
 CC EMBL; AF071946; AAD43405.1; -;  
 CC EMBL; AF071936; AAD43405.1; JOINED.  
 CC EMBL; AF071937; AAD43405.1; JOINED.  
 CC EMBL; AF071938; AAD43405.1; JOINED.  
 CC EMBL; AF071939; AAD43405.1; JOINED.  
 CC EMBL; AF071940; AAD43405.1; JOINED.  
 CC EMBL; AF071941; AAD43405.1; JOINED.  
 CC EMBL; AF071942; AAD43405.1; JOINED.  
 CC EMBL; AF071943; AAD43405.1; JOINED.  
 CC EMBL; AF071944; AAD43405.1; JOINED.  
 CC EMBL; AF071945; AAD43405.1; JOINED.  
 CC EMBL; AF071946; AAD43406.1; -;  
 CC EMBL; AF071936; AAD43406.1; JOINED.  
 CC EMBL; AF071937; AAD43406.1; JOINED.  
 CC EMBL; AF071938; AAD43406.1; JOINED.  
 CC EMBL; AF071939; AAD43406.1; JOINED.  
 CC EMBL; AF071940; AAD43406.1; JOINED.  
 CC EMBL; AF071941; AAD43406.1; JOINED.  
 CC EMBL; AF071942; AAD43406.1; JOINED.  
 CC EMBL; AF071943; AAD43406.1; JOINED.  
 CC EMBL; AF071944; AAD43406.1; JOINED.  
 CC EMBL; AF071945; AAD43406.1; JOINED.  
 CC EMBL; AF071946; AAD43407.1; -;  
 CC EMBL; AF071936; AAD43407.1; JOINED.  
 CC EMBL; AF071937; AAD43407.1; JOINED.  
 CC EMBL; AF071938; AAD43407.1; JOINED.  
 CC EMBL; AF071939; AAD43407.1; JOINED.  
 CC EMBL; AF071940; AAD43407.1; JOINED.  
 CC EMBL; AF071941; AAD43407.1; JOINED.  
 CC EMBL; AF071942; AAD43407.1; JOINED.  
 CC EMBL; AF071943; AAD43407.1; JOINED.  
 CC EMBL; AF071944; AAD43407.1; JOINED.  
 CC EMBL; AF071945; AAD43407.1; JOINED.  
 CC EMBL; AF071946; AAD43407.1; JOINED.

DR EMBL; AF071944; AAD43407.1; JOINED.  
 DR EMBL; AF071945; AAD43407.1; JOINED.  
 DR EMBL; X55663; CAA39196.1; -;  
 DR EMBL; M33427; AAA40018.1; -;  
 DR PIR; JU0215; JU0215.  
 DR PIR; T01380; T01380.  
 DR PDB; 1GL5; 28-NOV-01.  
 DR MGD; MGI:98662; Tec.  
 DR InterPro; IPR001562; BTK.  
 DR InterPro; IPR001849; PH.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001452; SH3.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF00018; SH3; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR Pfam; PF00169; PH; 1.  
 DR Pfam; PF00779; BTK; 1.  
 DR PRINTS; PR00401; SH2DOMAIN.  
 DR PRINTS; PR00452; SH3DOMAIN.  
 DR PRINTS; PR00402; TECBTKDOMAIN.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR ProDom; PD000093; SH2; 1.  
 DR ProDom; PD000066; SH3; 1.  
 DR SMART; SM00107; BTK; 1.  
 DR SMART; SM00233; PH; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR SMART; SM00326; SH3; 1.  
 DR SMART; SM00219; TyRK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00001; SH2; 1.  
 DR PROSITE; PS00002; SH3; 1.  
 DR PROSITE; PS00003; PH\_DOMAIN; 1.  
 KW Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;  
 KW SH3 domain; phosphorylation; Alternative splicing; 3D-structure.  
 FT DOMAIN 4 111 PH.  
 FT DOMAIN 178 238 SH2.  
 FT DOMAIN 246 344 SH3.  
 FT DOMAIN 369 622 PROTEIN\_KINASE.  
 FT NP\_BIND 375 383 ATP (BY SIMILARITY).  
 FT BINDING 387 397 ATP (BY SIMILARITY).  
 FT ACT\_SITE 488 518 BY SIMILARITY.  
 FT MOD\_RES 518 518 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT VARSPLIC 1 94 MNFTILBEILIKRSQKKTKSLNLYKERLCVLPKSVLSY  
 EGRAKKYKRGVIDISKIKCVIKVNDGVIPOCKKFFQV  
 VHDANTLYIFAP -> MVVSFFVKINFHS (in isoform  
 2).  
 FT FTId=VSP\_005012.  
 VVDANTLYIFAPSPQSRD -> STKQSPMGSEVKRNKEQ  
 Q (in isoform 6).  
 FT FTId=VSP\_005013.  
 Missing (in isoform 6).  
 FT FTId=VSP\_005014.  
 Missing (in isoform 2, isoform 4 and  
 isoform 5).  
 FT FTId=VSP\_005015.  
 RPEGRPSLEILRTDELVECEETFGF -> ESCLCRVAQD  
 LSSKNLIGSRF (in isoform 3 and isoform 4).  
 FT FTId=VSP\_005016.  
 L -> P (IN REF. 2).  
 C -> F (IN REF. 2).  
 P -> T (IN REF. 2).  
 V -> E (IN REF. 4).  
 FGVL -> YGIP (IN REF. 4).  
 T -> S (IN REF. 2).  
 L -> F (IN REF. 2 AND 3).  
 SQ SEQUENCE 630 AA; 73426 MW; 262640EE90D4A6D2 CRC64;

Alignment Scores:



```

Pred. No.: 6,54e-22 Length: 630
Score: 301.50 Matches: 71
Percent Similarity: 52.51% Conservative: 44
Best Local Similarity: 32.42% Mismatches: 89
Query Match: 24.08% Indels: 15
DB: 1 Gaps: 7

US-09-830-144-1_COPY_408_1091 (1-684) x TEC_MOUSE (1-630)

QY 4 GAGCTTCGGAGCTATCCCGTGAACCATCTAATATGTTAAAGCTTTTGGAGCCTGC 63
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 412 GluAlaLysValMetMetLysLeuThrHisProLysLeuValGlnLeuTyrGlyValCys 431
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 64 TTG-----AATCCAGTGTCTCTGTCATGGAATATCTGAAGGGGCTCTTATATAAT 117
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 432 ThrGlnGlnLysProIleTyrIleValThrGluPheMetGluArgGlyCysLeuAsn 451
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 118 GTGCTGTCATGCTGTGAACCATTCGCCATATATATCTGCTGCCACGCAATGAGTTGGTGT 177
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 452 PheLeuArgGlnArgGln-----GlyHisPheSerArgAspMetLeuLeuSerMetCys 469
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 178 TTACAGTGTTCACAGAGTGGCTTATCTTACAGCATGCAACCAAGCGCTAATTCAC 237
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 470 GlnAspValCysGluGlyMetGluTyrLeu-----GluArgAsnSerPheIleHis 486
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 238 AGGACCTGAACACCAACCACTTACTGCTGCTGCAGGGGGACAGTCTTAAAAATTGT 297
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 487 ArgAspLeuAlaAlaArgAsnCysLeuValAsnGluAlaGly----ValValLysValSer 505
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 298 GATTTTGGTACAGCC-----TGTGACATTCAGACACACATGACCAATAACAAG 345
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 506 AspPheGlyMetAlaArgTyrValLeuAspAspGlnTyrThrSerSerGlyValAlaLys 525
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 346 GGGAGTCTGCTGGAGTGGACCTGGAAGTTTTCAGAGTAGTAATACAGTGAAATATGT 405
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 526 PheProValLysTrpCysProGluValPheAsnTyrSerArgPheSerSerLysSer 545
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 406 GAGCTCTTCAGTGGGTATATCTTTGGAGGTGATAGCGCT---CGAAGACCCCTTT 462
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 546 AspValTrpSerPheGlyValLeuMetTrpGluIlePheThrGluGlyArgMetProPhe 565
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 463 GATGAGATTGTGGCCCGAGCTTTCGGAATCATGTGGGTGTTTCATATGGTACTCGACCA 522
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 566 GluLys-----AsnThrAsnTyrGluValValThrMetValThrArgGlyHisArgLeu 583
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 523 CCACGTGATAAAAAATTTACCTAAGCCCATTCAGAGCCCTGATGACTCGTGTGTGCTCTAAA 582
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 584 HisArgProLysLeuAlaThrLysTyrLeuTyrGluValMetLeuArgCysTrpGlnGlu 603
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 583 GATCCTTCCAGCGCCCTTCATGAGGAGAAATTTGTGAAATAATGACTCACTTGTATG 639
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 604 ArgProGluGlyArgProSerLeuGluAspLeuLeuThrIleAspGluLeuVal 622

RESULT 14
TEC_HUMAN
ID_1 TEC_HUMAN STANDARD; PRT; 631 AA.
AC P42680;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase Tec (EC 2.7.1.112).
GN TEC OR PSTK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95019807; PubMed=7934162;
RA Sato K., Mano H., Ariyama T., Inazawa J., Yazaki Y., Hirai H.;
RT "Molecular cloning and analysis of the human Tec protein-tyrosine
kinase."

```

Alignment Scores:



THYMOCYTES THAN IN PERIPHERAL T CELLS, AND INCREASE IN THE THYMUS DURING DEVELOPMENT FROM NEONATE TO ADULT.

-1- INDUCTION: By interleukin-2.

-1- DOMAIN: THE PH DOMAIN MEDIATES TARGETING OF ITK/TSK AND IS INDISPENSIBLE FOR THE ACTIVATION THROUGH TCR/CD3.

-1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. TEC SUBFAMILY.

-1- SIMILARITY: Contains 1 SH2 domain.

-1- SIMILARITY: Contains 1 SH3 domain.

-1- SIMILARITY: Contains 1 PH domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; L00619; AAA39337.1; -;  
 CC EMBL; L05631; AAA40518.1; -;  
 CC EMBL; L10628; -; NOT\_ANNOTATED\_CDS.  
 CC EMBL; D14042; BAA03129.1; -;  
 CC PIR; A43030; A43030.  
 CC PDB; 1AWJ; 14-JAN-98.  
 CC PDB; 1LUK; 27-NOV-02.  
 CC PDB; 1LUM; 27-NOV-02.  
 CC PDB; 1LUN; 27-NOV-02.  
 CC MGD; MGI:96621; Itk.  
 CC InterPro; IPR001562; BTK.  
 CC InterPro; IPR001849; PH.  
 CC InterPro; IPR000719; Prot\_kinase.  
 CC InterPro; IPR000980; SH2.  
 CC InterPro; IPR001452; SH3.  
 CC InterPro; IPR001245; Tyr\_pkinase.  
 CC Pfam; PF00779; BTK; 1.  
 CC Pfam; PF00169; PH; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC Pfam; PF00017; SH2; 1.  
 CC Pfam; PF00018; SH3; 1.  
 CC PRINTS; PR00401; SH2DOMAIN.  
 CC PRINTS; PR00402; TBCBTDOMAIN.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC ProDom; PD000001; Prot\_kinase; 1.  
 CC ProDom; PD000093; SH2; 1.  
 CC ProDom; PD000066; SH3; 1.  
 CC SMART; SM00107; BTK; 1.  
 CC SMART; SM00233; PH; 1.  
 CC SMART; SM00252; SH2; 1.  
 CC SMART; SM00326; SH3; 1.  
 CC SMART; SM00219; TyzKc; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE; PS50001; SH2; 1.  
 CC PROSITE; PS50002; SH3; 1.  
 CC PROSITE; PS50003; PH\_DOMAIN; 1.  
 CC Transferase; Tyrosine-protein kinase; Phosphorylation;  
 CC ATP-binding; SH2 domain; SH3 domain; 3D-structure.  
 CC DOMAIN 4 117  
 CC PH.  
 CC DOMAIN 177 237  
 CC SH2.  
 CC DOMAIN 245 343  
 CC SH2.  
 CC DOMAIN 368 620  
 CC PROTEIN\_KINASE.  
 CC NP\_BIND 374 382  
 CC ATP (BY SIMILARITY).  
 CC BINDING 396 396  
 CC ATP (BY SIMILARITY).  
 CC ACT\_SITE 487 487  
 CC BY SIMILARITY.  
 CC MOD\_RES 517 517  
 CC PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 CC CONFLICT 82 87  
 CC MISSING (IN REF. 2, 3 AND 4).  
 CC CONFLICT 535 535  
 CC F -> S (IN REF. 3).  
 CC CONFLICT 540 540  
 CC STRAND 188 188  
 CC TURN 193 194  
 CC STRAND 198 198

FT TURN 209 210  
 FT STRAND 214 214  
 FT STRAND 228 230  
 FT TURN 230 231  
 SQ SEQUENCE 625 AA; 72291 MW; F7A4A18A8A1AADDCC CRC64;

#### Alignment Scores:

Pred. No.: 1.03e-21 Length: 625  
 Score: 299.50 Matches: 72  
 Percent Similarity: 48.66% Conservative: 37  
 Best Local Similarity: 32.14% Mismatches: 88  
 Query Match: 23.92% Indels: 27  
 DB: 1 Gaps: 9

US-09-830-144-1\_COPY\_408\_1091 (1-684) x ITK\_MOUSE (1-625)

Qy 4 GAGCTTCGGCAGGTATCCCGTGTGAACCATCTTAATATGTAAGCTTTATGAGCCTTCG 63  
 Db 411 GluAlaGluValMetMetLysLeuSerHisProLysLeuValGlnLeuTyrGlyValCys 430  
 Qy 64 TTGAAT-----CCAGTGTGCTTGTGATGAATATGCTGAAGGGGGCTCTTTATATAAT 117  
 Db 431 LeuGluGlnAlaProLleCysLeuValPheGluPheMetGluHisGlyCysLeuSerAsp 450  
 Qy 118 GTGCTGCATGGTGTGCTGAACCATCTTACTCATATATATCTGCTGCCACGCAATGAGTTGGT 177  
 Db 451 TyrLeuArgSerGlnArgGlyLeu-----PheAlaAlaGluThrLeuLeuGlyMetCys 468  
 Qy 178 TTACAGTGTTCCTCCCAAGGAGTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCAC 237  
 Db 469 LeuAspValCysGluGlyMetAlaTyrLeu-----GluLysAlaCysValIleHis 485  
 Qy 238 AGGACCTGAACACCAACTTACTCTGCTGAGGGGGGAGCAGTCTTAAAAATTTGT 297  
 Db 486 ArgAspLeuAlaAlaArgAsn---CysLeuValGlyGluAsnGlnValIleLysValSer 504  
 Qy 298 GATTTTGGT-----ACAGCCTGTGCACATTCAGACACACATGACCAATAACAAG 345  
 Db 505 AspPheGlyMetThrArgPheValLeuAspGlnTyrThrSerSerThrGlyThrLys 524  
 Qy 346 GGGAGTCTGCTTGGATGGCAGCTGAAGTTTTCAGGTAGTATTAATACAGTGAAAAATGT 405  
 Db 525 PheProValLysTyrAlaSerProGluValPheSerPheSerArgTyrSerSerLysSer 544  
 Qy 406 GAGCTCTTCAGCTGGGTATTATTCTTTGGGAAGTATAACGGCTCGGAAA---CCCTTT 462  
 Db 545 AspValTyrSerPheGlyValLeuMetTyrGluValPheSerGluGlyLysIleProTyr 564  
 Qy 463 GATCAGATTGGTGGCCCGAGCTTCCGAAATCATGTGGGCTGTTCAATATGGTACTCGACCA 522  
 Db 565 GluAsnArgSerAsnSer-----GluValValGluAspIleSerThrGlyPheArg--- 581  
 Qy 523 CCAGTGATAAAAAATTACCTTAAGCC-----ATTGAGAGCTCGATG 564  
 Db 582 -----LeuTyrLysProArgLeuAlaSerCysHisValTyrGlnIleMet 596  
 Qy 565 ACTGTTGTTGGTCTAAAGATCTTCACAGCGCCCTTCAATGGAGGAAATGTGAAATA 624  
 Db 597 AsnHisCysTyrLysGluLysProGluAspArgProPheSerGlnLeuLeuSerGln 616  
 Qy 625 ATGACTCACTTG 636  
 Db 617 LeuAlaGluIle 620

Search completed: December 4, 2003, 09:03:39

Job time : 25.1757 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 08:22:54 ; Search time 77.7973 Seconds  
(without alignments)  
4537.638 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091

Perfect score: 1252

Sequence: 1 gtgagcttcggcagttatc.....cattacagtatcctgttcag 684

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09830144/runat\_03122003\_122344\_21277/app\_query.fasta\_1.1230  
-DB=SPTRMBL\_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144\_@CGN\_1\_100 @runat\_03122003\_122344\_21277 -NCPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTRMBL\_23:\*\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rviro:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1252	100.0	478 4 Q9NTR4	Q9ntr4 homo sapien

ID	Q9NTR4	PRELIMINARY;	PRT;	478 AA.
2	1252	100.0	491	4 Q9NZ70
3	1252	100.0	539	4 Q9NTR1
4	1252	100.0	566	4 Q9NTR2
5	1252	100.0	606	11 Q923A8
6	1221	97.5	616	13 Q73613
7	700	55.9	678	5 Q9V3Q6
8	414	33.1	258	5 Q9VCV0
9	379	30.3	422	10 Q8GV29
10	376	30.0	371	13 Q90ZY8
11	375.5	30.0	2631	5 Q8MYR1
12	374.5	29.9	2964	5 Q8I7W7
13	374	29.9	411	10 Q9ZQ31
14	373	29.8	412	10 Q9M085
15	373	29.8	417	10 Q8GV30
16	372	29.7	289	11 Q8BR73
17	372	29.7	454	11 Q9ESL3
18	372	29.7	802	11 Q9ESL4
19	371	29.6	455	4 Q9HCC4
20	371	29.6	800	4 Q9NVE9
21	371	29.6	800	4 Q9NVE9
22	371	29.6	800	4 Q9NVE9
23	371	29.6	800	4 Q9HDD2
24	368	29.4	416	10 Q94C42
25	366	29.2	637	10 Q94J41
26	363.5	29.0	1148	5 Q95VF6
27	363.5	29.0	1161	5 Q95UN8
28	362	28.9	422	5 Q23846
29	362	28.9	1338	5 Q23927
30	360.5	28.8	608	11 Q8BIG8
31	360.5	28.8	1066	4 Q9H2N5
32	356.5	28.5	888	11 Q8CEX3
33	356	28.4	462	10 Q39886
34	355.5	28.4	370	10 Q9S7D5
35	355	28.4	421	10 Q8GV28
36	355	28.4	564	4 Q9H1Y7
37	354.5	28.3	1161	5 Q8MEK7
38	353.5	28.2	859	4 Q8WY25
39	353.5	28.2	888	11 Q8CDL6
40	351.5	28.1	977	5 Q9VMW24
41	351	28.0	570	4 Q8WVN2
42	351	28.0	1036	4 Q8WVN1
43	350	28.0	546	10 Q22558
44	349	27.9	1001	11 Q8VDG6
45	348	27.8	391	10 Q8LCP3

#### ALIGNMENTS

RESULT 1

Q9NTR4 ID Q9NTR4 PRELIMINARY; PRT; 478 AA.  
AC Q9NTR4;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE D154G14.1.3 (Mitogen-activated protein kinase kinase kinase 7 (TGF-beta activated kinase 1c (TAK1))) (Fragment).  
GN MAP3K7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP Tracey A.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AL121964; CAB87604.1; -.  
DR HSSP; P12931; 1FMK.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser thr kinase.  
DR InterPro; IPR001245; Tyr\_kinase.  
DR Pfam; PF00069; pkinase; 1.

DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
FT NON\_TER 1  
SQ SEQUENCE 478 AA; 52482 MW; 177CC8CFA8D8B8F8 CRC64;  
  
Alignment Scores:  
Pred. No.: 5,79e-134 Length: 478  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-830-144-1\_COPY\_408\_1091 (1-684) x Q9NZ70 (1-478)  
  
QY 1 GTAGAGCTTCGGCAGTATCCGCTGTGAACCTCCTAAATATTCTAAAGCTTTATGAGCC 60  
Db ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 55  
QY 61 TGCCTTGAATCCAGTGTCTTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
Db CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 75  
QY 121 CTGATGCTGCTGAACCATTCGATGATATATCTGCTGCCACGCAATGATGTTGTTTA 180  
Db LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 95  
QY 181 CAGTGTTCACAGGAGTGTCTATCTTACAGCATCAACCAACGCGTAAATTCACAG 240  
Db GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 115  
QY 241 GACCTGAAACCCAACTTACTGCTGTGAGGGGGACAGTCTTAAATTTTGTGAT 300  
Db AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 135  
QY 301 TTTGTTACAGCTGTGACATTCAGACACATGACCAATTAACAAGGGAGTGCTGTTGG 360  
Db PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 155  
QY 361 ATGGCAGCTGAAGTGTCTTGAAGTACTAATTCAGTGAAGGAGTCTTCACTCTCAGCTGG 420  
Db MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 175  
QY 421 GGTATTATTCTTTGGGAAGTGAACGCTGCGAAACCCCTTTGATGAGATTTGTTGCCCA 480  
Db GlyIleLeuLeuTrpGluValIleThrArgGlySerProPheAspGluIleGlyGlyPro 195  
QY 481 GCTTTCCGAATCATGCGGTGTTTCAATATGTTACTCGACACCACTGATATAAAATTTA 540  
Db AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 215  
QY 541 CCTAACCCATTGAGAGCTGATGCTGCTGTGCTTAAAGATCCTTCCAGCGCCCT 600  
Db ProLysProIleGlnSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 235  
QY 601 TCAATGAGGAAATTTGAAATATATGACTCATCTGATCGGTACTTCCAGGAGCAGAT 660  
Db SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 255  
QY 661 GAGCAATTACAGTATCCTTGTGTCAG 684  
Db GluProLeuGlnTyrProCysGln 263

RESULT 2

Q9NZ70

ID Q9NZ70 PRELIMINARY; PRT; 491 AA.

AC Q9NZ70; Q9NZ70;

DT 01-OCT-2000 (T-EMBLrel. 15, Created)

DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)

DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)  
DE TGF beta-activated kinase splice variant d (DJ154G14.1.4) (Mitogen-  
DE activated protein kinase kinase kinase 7 (TGF-beta activated kinase 1d  
DE (TAK1)))  
GN TAK1 OR MAP3K7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20568765; PubMed=11118615;  
RA Dempsey C.E., Sakurai H., Sugita T., Guesdon F.;  
RT "Alternative splicing and gene structure of the transforming growth  
RT factor beta-activated kinase 1."; (2000).  
RL Biochim. Biophys. Acta 1517:46-52 (2000).  
RN [2]  
RP SEQUENCE OF 41-491 FROM N.A.  
RA Tracey A.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AF218074; AAF27652.1; -.  
DR EMBL; AL121964; CAB87605.1; -.  
DR HSSP; P08631; IAD5.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser Thr kinase.  
DR InterPro; IPR001245; Tyrosine kinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 491 AA; 53739 MW; B7D832E286A99C5 CRC64;  
  
Alignment Scores:  
Pred. No.: 5,83e-134 Length: 491  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-830-144-1\_COPY\_408\_1091 (1-684) x Q9NZ70 (1-491)  
  
QY 1 GTAGAGCTTCGGCAGTATCCGCTGTGAACCTCCTAAATATTCTAAAGCTTTATGAGCC 60  
Db ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 95  
QY 61 TGCCTTGAATCCAGTGTCTTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
Db CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 115  
QY 121 CTGATGCTGCTGAACCATTCGATGATATATCTGCTGCCACCAATGATGTTGCTTTTA 180  
Db LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 135  
QY 181 CAGTGTTCACAGGAGTGTCTTATCTTACAGCATCAACCAACGCGTAAATTCACAG 240  
Db GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 155  
QY 241 GACCTGAAACCCAACTTACTGCTGTGAGGGGGACAGTCTTAAATTTTGTGAT 300  
Db AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 175  
QY 301 TTTGTTACAGCTGTGACATTCAGACACATGACCAATTAACAAGGGAGTGCTGTTGG 360  
Db PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195  
QY 361 ATGGCAGCTGAAGTGTCTTGAAGTACTAATTCAGTGAAGGAGTCTTCACTCTCAGCTGG 420  
Db MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215

QY 421 GGTATTATTCTTTGGGAAGTGAACCGCTCGGAACCCCTTTGATGAGATTGGTGCCCA 480  
 DB 216 GlylleleleuTrpGluValleThrArgArgLysPropheAspGluilleGlyPro 235  
 QY 481 GCTTTCCGAATCATGTGGCTGTTTCATATGCTACTCGACACCACTGATAAAAAATTTA 540  
 DB 236 AlaPheArgilleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 255  
 QY 541 CCTAAGCCCATGAGAGCTGATGACTCTGTTGGTCTTAAGATCCTTCCAGGCGCCT 600  
 DB 256 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 275  
 QY 601 TCAATGAGAGAAATGTGAAATAATGACTCACTTGTGCGGTACTTTCAGGAGCAGAT 660  
 DB 276 SerMetGluGluIleValleLysleMetThrHisLeuMetArgTrpPheProGlyAlaAsp 295  
 QY 661 GAGCATTACAGTATCTCTGTCAG 684  
 DB 296 GluProLeuGlnTyProCysGln 303

## RESULT 3

Q9NTR1 ID Q9NTR1 PRELIMINARY; PRT; 539 AA.  
 AC Q9NTR1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE D154G14.1.1 (Mitogen-activated protein kinase kinase 7 (TGF-  
 beta activated kinase 1a (TAK1))) (Fragment).  
 GN MAP3K7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tracey A.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AL121964; CAB87607.1; -.  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 FT NON\_TER 1  
 SQ SEQUENCE 539 AA; 60006 MW; E6183F553CC7F324 CRC64;

## Alignment Scores:

Pred. No.: 5,97e-134 Length: 539  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x Q9NTR1 (1-539)

QY 1 GTAGAGCTCGCAGTATCCCGTGTGAACCATCTATATGTAAAGCTTTATGAGCC 60  
 DB 36 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyGlyAla 55  
 QY 61 TGCTTGAATCCAGTGTCTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
 DB 56 CysLeuAsnProValCysLeuValMetGluTyAlaGluGlySerLeuTyAsnVal 75

QY 121 CTGCATGTGTCTGACCACTTGCATATTATATCTGTCGCCAGCAATGAGTTGGTGTTA 180  
 DB 76 LeuHisGlyAlaGluProLeuProTyTrpThrAlaAlaHisAlaMetSerTrpCysLeu 95  
 QY 181 CAGTGTTCCTCCAGGAGTGGCTTACTTCCACAGCATGCAACCCAAAGCCCTAATTCACAGG 240  
 DB 96 GlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProLysAlaLeuIleHisArg 115  
 QY 241 GACCTGAAACCAACCAACTTACTGCTGCTGCTGAGGGGCAAGTCTTAAAAATTTGTGAT 300  
 DB 116 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 135  
 QY 301 TTTGCTACAGCTGTGACCTTCAGACACATGACCATTAACAAGGGAGTCTCTCTGG 360  
 DB 136 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 155  
 QY 361 ATGCACCTGAAGTTTTTGAAGGTAGTAATPACAGTGAATAATGTGACGCTCTTCAGCTGG 420  
 DB 156 MetalaProGluValPheGluGlySerAsnTySerGluLysCysAspValPheSerTrp 175  
 QY 421 GGTATTATTCTTTGGGAAGTGAACCGCTCGGAACCCCTTTGATGAGATTGGTGCCCA 480  
 DB 176 GlylleleleuTrpGluValleThrArgArgLysPropheAspGluilleGlyGlyPro 195  
 QY 481 GCTTTCCGAATCATGTGGCTGTTTCATATGCTACTCGACACCACTGATAAAAAATTTA 540  
 DB 196 AlaPheArgilleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 215  
 QY 541 CCTAAGCCCATGAGAGCTGATGACTCTGTTGGTCTTAAGATCCTTCCAGGCGCCT 600  
 DB 216 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 235  
 QY 601 TCAATGAGAGAAATGTGAAATAATGACTCACTTGTGCGGTACTTTCAGGAGCAGAT 660  
 DB 236 SerMetGluGluIleValleLysleMetThrHisLeuMetArgTrpPheProGlyAlaAsp 255  
 QY 661 GAGCATTACAGTATCTCTGTCAG 684  
 DB 256 GluProLeuGlnTyProCysGln 263

## RESULT 4

Q9NTR2 ID Q9NTR2 PRELIMINARY; PRT; 566 AA.  
 AC Q9NTR2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE D154G14.1.2 (Mitogen-activated protein kinase kinase 7 (TGF-  
 beta activated kinase 1b (TAK1))) (Fragment).  
 GN MAP3K7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tracey A.;  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; AL121964; CAB87606.1; -.  
 DR HSSP; P12931; 1FMK.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 FT NON\_TER 1  
 SQ SEQUENCE 566 AA; 62972 MW; 32CDAC1211B200CF CRC64;

Alignment Scores:  
 Pred. No.: 6,04e-134 Length: 566  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x Q9NTR2 (1-566)

QY	1	GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATGTAAGCTTTATGAGCC	60
Db	36	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla	55
QY	61	TGCTGAATCCAGTGTCTCTGTGATGGAATATGTCGAAGGGGCTTTATATAATG	120
Db	56	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal	75
QY	121	CTGCATGCTGGAACCATGTCATATTAATCTGCTGCCCGCATGAGTTGCTGTTA	180
Db	76	LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu	95
QY	181	CAGTGTTCCTCAAGGAGTGTCTCTGATGGAATATGTCGAAGGGGCTTTATATAATG	240
Db	96	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg	115
QY	241	GACCTGAAACCCACCAATCTACTGCTGTTGCGGGGGACAGTCTTAAATAATTTGAT	300
Db	116	AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp	135
QY	301	TTTGGTACAGCTGTGACATTCAGACACATGCAACCAACCAAGGGAGTGTGCTGG	360
Db	136	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr	155
QY	361	ATGCGACCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	420
Db	156	MetAlaProGluValPheGluGlySerAsnTyrSerGlnCysAspValPheSerTyr	175
QY	421	GCTTTCCGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
Db	176	GlyIleLeuLeuTyrGluValIleThrArgLysProPheAspGluIleGlyGlyPro	195
QY	481	GCTTTCCGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540
Db	196	AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProLeuLeuLysAsnLeu	215
QY	541	CCTAAGCGAATTTGAAATAATGATGATGATGATGATGATGATGATGATGATGATGATG	600
Db	216	ProLysProIleLeuTyrGluValIleThrArgLysProPheAspGluIleGlyGlyPro	235
QY	601	TCAATGAGGAATTTGAAATAATGATGATGATGATGATGATGATGATGATGATGATGATG	660
Db	236	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp	255
QY	661	GAGCCATTACAGTATCTTCTGTCAG	684
Db	256	GluProLeuGlnTyrProCysGln	263

RESULT 5  
 Q923A8 PRELIMINARY; PRT; 606 AA.

AC Q923A8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Unknown (Protein for MGC:5989).  
 GN MAP3K7.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL; BC006665; AA06665.1; -.  
 DR MGI; MGI:1346877; Map3k7.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00669; pkinase; 1.  
 DR PRINTS; PR0109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 606 AA; 67194 MW; AB8664F389272102 CRC64;

Alignment Scores:  
 Pred. No.: 6,15e-134 Length: 606  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0

US-09-830-144-1\_COPY\_408\_1091 (1-684) x Q923A8 (1-606)

QY	1	GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATGTAAGCTTTATGAGCC	60
Db	76	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla	95
QY	61	TGCTGAATCCAGTGTCTCTGTGATGGAATATGTCGAAGGGGCTTTATATAATG	120
Db	96	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal	115
QY	121	CTGCATGCTGGAACCATGTCATATTAATCTGCTGCCCGCATGAGTTGCTGTTA	180
Db	116	LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu	135
QY	181	CAGTGTTCCTCAAGGAGTGTCTCTGATGGAATATGTCGAAGGGGCTTTATATAATG	240
Db	136	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg	155
QY	241	GACCTGAAACCCACCAATCTACTGCTGTTGCGGGGGACAGTCTTAAATAATTTGAT	300
Db	156	AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp	175
QY	301	TTTGGTACAGCTGTGACATTCAGACACATGCAACCAACCAAGGGAGTGTGCTGG	360
Db	176	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr	195
QY	361	ATGCGACCTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	420
Db	196	MetAlaProGluValPheGluGlySerAsnTyrSerGlnCysAspValPheSerTyr	215
QY	421	GCTTTCCGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	480
Db	216	GlyIleLeuLeuTyrGluValIleThrArgLysProPheAspGluIleGlyGlyPro	235
QY	481	GCTTTCCGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540
Db	236	AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProLeuLeuLysAsnLeu	255
QY	541	CCTAAGCGAATTTGAAATAATGATGATGATGATGATGATGATGATGATGATGATGATG	600
Db	256	ProLysProIleLeuTyrGluValIleThrArgLysProPheAspGluIleGlyGlyPro	275
QY	601	TCAATGAGGAATTTGAAATAATGATGATGATGATGATGATGATGATGATGATGATGATG	660
Db	276	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp	295
QY	661	GAGCCATTACAGTATCTTCTGTCAG	684



```

Db      296 GluProLeuGlnTyrProCysGln 303
|||||
RESULT 6
073613 PRELIMINARY; PRT; 616 AA.
ID 073613
AC 073613;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE TAK1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=98130593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAK1 and TAB1 in BMP signaling in early Xenopus
RT development.";
RL EMEQ J. 17:1019-1028 (1998).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U92030; AAC14008.1; --
DR HSP; P12931; 1FMK.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_Thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 616 AA; 68464 MW; 493AD2A05ADC38B6 CRC64;

Alignment Scores:
Pred. No.: 2,16e-130 Length: 616
Score: 1221.00 Matches: 222
Percent Similarity: 98.68% Conservative: 3
Best Local Similarity: 97.37% Mismatches: 3
Query Match: 97.52% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-1_COPY_408_1091 (1-684) x 073613 (1-616)
QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCCTTAATTTGTAAGCTTTATGGAGCC 60
Db 65 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuVallyLeuTyrGlyAla 84
QY 61 TGCTGAATCCAGTGTCTGTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120
Db 85 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 104
QY 121 CTGCATGCTGCTGAACCATGCTCATATATATCTGTCGCCACCAATGAGTGGTGTTRA 180
Db 105 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 124
QY 181 CAGTGTTCACCAAGAGTGGCTTATCTTCACAGCATGCACACCAAGCGCTTAATTCACAG 240
Db 125 GlnCysAlaGlnGlyValAlaTyrLeuHisSerMetCysProlyAlaLeuHisArg 144
QY 241 GACCTGAACACCAACCAACTTACTGCTGGTGTGAGGGGACAGTCTTAAATAATTCGTAT 300
Db 145 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysLeuCysAsp 164
QY 301 TTGGTACAGCTGTGACATTCAGACACATGACCAATACCAAGGGAGTGTGCTGTGG 360
Db 165 PheGlyThrAlaCysAspIleGlnThrHisMetAsnAsnLysGlySerAlaAlaTTP 184

```

```

QY 361 ATGGCAGCTGAAGCTTTTGAAGGTAGTAATAATACAGTGAATAATGAGCTTTCAGCTGG 420
Db 185 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 204
QY 421 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAAACCCCTTTGATGATGATGGTGGCCCA 480
Db 205 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 224
QY 481 GCTTTCGAATCATGTGGGTGTTCATAATGTTACTCGACACCACTGATATAAAAAATTTA 540
Db 225 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 244
QY 541 CCTAAGCCCATTTGAGAGCCTGATGACATCGTGTGGTCTAAGATCCTTCCACGGCCT 600
Db 245 ProlysProIleGluSerLeuMetThrArgCysTrpSerLysAspProProGlnArgPro 264
QY 601 TCAATGGAGGAAATTTGTGAAATAATGACTCAGTTCATGTCGGTACTTTCAGGAGCAGAT 660
Db 265 SerMetGluGluIleValIleMetThrHisLeuMetGlnTyrPheProGlyAlaAsp 284
QY 661 GAGCCATTACAGTATCCTTGTCTAG 684
Db 285 ValSerLeuGlnTyrProCysGln 292

RESULT 7
QY9V3Q6 PRELIMINARY; PRT; 678 AA.
ID QY9V3Q6
AC QY9V3Q6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative TAK1 protein (CG1386 protein) (L042274P).
GN TAK1 OR CG1386 OR CG18492.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

```

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*,"  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Takatsu Y., Nakamura M., Stapleton M., Danos M., Matsumoto M.,  
RA O'Connor M.B., Shibuya H., Ueno N.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AF003571; AAF50895.1; -;  
DR EMBL; AF199466; AAF06815.1; -;  
DR EMBL; AY051953; AAK93377.1; -;  
DR HSP; P08631; IAD5;  
DR Flybase; FBGN026323; Tak1.  
DR InterPro; IPR001990; Granin.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser Thr pkinase.  
DR InterPro; IPR001245; Tyr pkinase.  
DR Pfam; PF00669; pkinase; 1.  
DR PRINTS; P00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR PROSITE; PS00422; GRANINS 1; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
KW Hypothetical protein; ATP-binding; Kinase;  
KW Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 678 AA; 75674 MW; 87EBA80CDB8CDE45 CRC64;

Alignment Scores:  
Pred. No.: 8, 2e-71  
Score: 700.00  
Percent Similarity: 73.01%  
Best Local Similarity: 57.52%  
Query Match: 55.91%  
Indels: 4  
Gaps: 3  
DB: 5

US-09-830-144-1\_COPY\_408\_1091 (1-684) x Q9V3Q6 (1-678)

QY 4 GAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATTTAGCTTATGA----- 57  
Db 60 GluVallysglnLeuSerArgVallyshisProAsnlelleleAlaLeuHisGlyIleSer 79  
QY 58 GCGTGTGATCCAGTGTCTGTGTCATGATATGCTCAAGGGGCTTTATATAAT 117  
Db 80 SerTyrGlnGlnAlaThrTyrlleuMetGluPheAlaGluGlySerLeuHisAsn 99  
QY 118 GTGCTGCATGTGTGATGAACCATGTCCTATATATCTCTGCCCAAGCAATGAGTTGTGT 177  
Db 100 PheLeuHisGly--LysVallyshisProAlaTyrSerLeuAlaHisAlaMetSerTrpAla 118  
QY 178 TTAAGTGTTCACAGAGTGGCTTATCTTACAGCATGCAACCAAGCGCTTAATTCAC 237  
Db 119 ArgGlnCysAlaGluGlyLeuAlaTyrleuHisAlaMetThrProLysProLeuIleHis 138  
QY 238 AGGAGCTGAACACCAACCACTTACTGTGTGTCACGGGGGACAGTTCTAAATAATTTGT 297  
Db 139 ArgaspVallyshisProLeuAsnleuLeuThrAsnLysGlyArgAsnleuIleCys 158  
QY 298 GATTTTGTACAGCTGTGACATTCAGACACACATGACCAATACAAAGGGAGTGTCT 357

Db 159 AspPheGlyThrValAlaAspLysSerThrMetMetThrAsnAsnArgGlySerAlaAla 178  
QY 358 TCGATGGCAGCTGAAGTATTTGAAGCTAGTATTAACAGTGAATAAATGTCAGCTTCTCAGC 417  
Db 179 TrpMetAlaProGluValPheGluGlySerLysTyrThrGluLysCysAspIlePheSer 198  
QY 418 TGGGGTATTATTCTTTGGGAAGTGATACGCGTCGGAACCCCTTTGATGAGATTGGTGGC 477  
Db 199 TrpAlaIleValLeuTrpGluValLeuSerArgLysGlnProPheLysGlyIleAspAsn 218  
QY 478 CCAGCTTTCCGAATCATGTGGCTGTTCTAATAGTGTCTCGACCACTGATAAAAT 537  
Db 219 ---AlaTyrThrIleGlnTrpLysIleTyrLysGlyGluArgProLeuLeuThrThr 237  
QY 538 TTACCTTAAGCCCATGAGAGCGCTGATGACTCGTGTGTGTCTTAAAGATCCTTCCCAGCGC 597  
Db 238 CysProLysArgIleGluLeuLeuMetThrAlaCysTrpLysThrValProGluAspArg 257  
QY 598 CCTTCAATGGAGGAAATTTGTGAATAATGACTCACTGATGCGGTACTTTCAGGAGCA 657  
Db 258 ProSerMetGlnTyrIleValGlyValMetHisGluIleValLysAspTyrThrGlyAla 277  
QY 658 GATGAGCCATTACAGTAT 675  
Db 278 AspLysAlaLeuGluTyr 283

RESULT 8  
Q9V3Q6  
ID Q9V3Q6 PRELIMINARY; PRT; 258 AA.  
AC Q9V3Q6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE CG4803-PA.  
GN TAKL2.  
OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; *Drosophila*.  
OX NCBI\_TaxID=7221;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman J.B., Bhandari D., Bolshakov S.,  
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Heston D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,  
RA Hsiao L., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Moberg C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Mount D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

Db	113	HisGlnIleAlaGlnGlyLeuHisGlyMetGlnProLysAlaValIleHis	133
Qy	238	AGGGACCTGAACACCACTTACTGCTGGTTGCAGGGGGACAGTTCTAAAAATTTGT	297
Db	133	ArgAspIleIleProLeuAsnThrLeuLeuLeuGlyLeuLysLeuLysIleCys	152
Qy	298	GATTTTGGTACAGCGCTGCACATTCAGACACACATGACCAATAACAAGGGGAGTGTGCT	357
Db	153	AspPheGlyThrValValAspLeuSerGlnSerIleSerCysAsnAlaGlyThrCysArg	172
Qy	358	TGGATGGCACTGAA	372
Db	173	TyrIysAlaProGluValArgGluLeuPheAspPheLysSerAsnArgIleIleAsn	192
Qy	373	-----GTTTTCGAAGGTAGTAATTACAGTGAATAATGCAAAATGTCACGTC	411
Db	193	GlnProThrGlyPheGlnLysValGlnGlyAsnLysProAspGluLysCysAspVal	212
Qy	412	TTACAGCTGGGGTATTATCTTTGGGAAGTGATAACGCTCGGAACCCCTTTTCATCAGATT	471
Db	213	TyrSerTrpAlaIleThrPheTrpGluIleLeuSerArgLysGluProPheGluGlnTyr	232
Qy	472	GGTGGCCCAAGCTTCCGAATCATGTGGCTGTTCAATATGGTACTCGA	519
Db	233	---AsnThrLeuPheGluLeuTyrMetAlaIleAsnGluGlyLysArg	247
RESULT 9			
Q8GV29	ID	Q8GV29	PRELIMINARY; PRT; 422 AA.
AC	Q8GV29;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DE	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Serine/threonine protein kinase.		
GN	DPK2.		
OS	Oryza sativa (japonica cultivar-group).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;		
OC	Ehrhartoideae; Oryzeae; Oryza.		
OX	NCBI_TaxID=39947;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RA	Gu Z.M., Zhang H.S., Huang J.;		
RT	"Molecular Cloning And Characterization Of A Novel Serine/Threonine		
RT	Protein Kinase Gene Family From Rice."		
RL	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY156511; AAN84503.1; -.		
KW	Kinase.		
SQ	SEQUENCE 422 AA; 46587 MW; BCC22F1FC6CFF62C CRC64;		
Alignment Scores:			
Pred. No.:	3.66e-34	Length:	422
Score:	379.00	Matches:	85
Percent Similarity:	58.18%	Conservative:	43
Best Local Similarity:	38.64%	Mismatches:	70
Query Match:	30.27%	Indels:	22
DB:	10	Gaps:	9
US-09-830-144-1_COPY_408_1091 (1-684) x Q8GV29 (1-422)			
Qy	4	GAGCTTCGGCAGTATATCCCGTGTGAACCATCTTAATTTAAAGCTTTATGAGCGCTGC	63
Db	190	GluValMetMetLeuAlaThrLeuArgHisProAsnIleValLysPheIleGlyAlaCys	209
Qy	64	TTGAATCCAGTG-----TGTCCTGTGATGGAATATGCTGAAGGGGCTCTTTATAAAT	117
Db	210	ArgLysProMetValTrpCysIleValThrGluTyrAlaLysGlyGlySerValArgGln	229
Qy	118	GTGCTG-----CATGCTGCTGAACCATTGCCATATTATATCTGCTGCCCGCA	165
Db	230	PheLeuMetLysArgGlnAsnArgSerValProLeuLysLeu-----Ala	244
Qy	166	ATGAGTTGGTGTTTACAGTGTCTCCCAAGAGAGTGGCTTATCTTCCACAGCAGTCAACCTAA	225

```

Db 245 VailysGlnAlaLeuAspValAlaArgGlyMetAlaTyrValHisAlaLeu----- 261
QY 226 GCGCTAATTCAGGAGCTGAACACCACTTACTGCTGGTTGCGAGGGGAGAGTT 285
Db 262 GlyPheIleHisArgAspLeuLysSerAspAan---LeuLeuIleSerGlyAspLysSer 280
QY 286 CTAATAATTTGATTTGGTACAGC---TGTGACATTCACACACAC---ATGACCAAT 339
Db 281 IleLysIleAlaAspPheGlyValAlaArgIleGluValLysThrGluGlyMetThrPro 300
QY 340 AACAGGGGAGGCTGCTGGATGCGACCTGAAGTTTGTGAAGGTAGTAATACAGTGAA 399
Db 301 GluThrGlyTyrArgTirPheAlaProGluMetIleGlnHisArgProTyrAspGln 320
QY 400 AAATGTGAGCTCTCAGCTGGGGTATTATCTTTGGGAAGTGAATACACGCGTGGAAACCC 459
Db 321 LysValAspValTyrSerPheGlyIleValLeuTyrPgluLeuIleThrGlyMetLeuPro 340
QY 460 TTTGATGAGATTGGTGGC-----CCAGCTTCCGAAATCATGTGGCTGTTCAATAGGT 513
Db 341 PheAlaAsnMetThrAlaValGlnAlaAlaPheAlaVal-----ValAsnLysGly 357
QY 514 ACTGACACCACTGATATAAATAATTTACCTAAGCCCATTCAGACCTGATGACTGCTTGT 573
Db 358 ValArgProAlaIleProGlnAspCysLeuProValLeuSerGluIleMetThrArgCys 377
QY 574 TGTCTAAAGATCTTCCAGCGCCCTTCAATGGAGGAATGTGAAATAATAGACTCAC 633
Db 378 TrpAspProAsnProAspValArgProPheThrGluValValArgMetLeuGluHis 397

RESULT 10
Q90ZY8 PRELIMINARY; PRT; 371 AA.
AC Q90ZY8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cyclic GMP-binding protein C.
GN GBPC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA Bosgraaf L., Russcher H., Smith J.L., Wessels D., Solls D.R.,
RA Van Haastert P.J.M.;
RT "A novel cGMP signalling pathway mediating myosin phosphorylation and
RT chemotaxis in Dictyostelium.";
RL EMBO J. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22008080; PubMed=12011437;
RA Goldberg J.M., Bosgraaf L., Van Haastert P.J., Smith J.L.;
RT "Identification of four candidate cGMP targets in Dictyostelium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:6749-6754(2002).
DR EMBL; AF481923; AAM34041.1; -.
DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR000591; DEP.
DR InterPro; IPR004182; GRAM dom.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000651; RasGEFN.

US-09-830-144-1_COPY_408_1091 (1-684) x Q90ZY8 (1-371)

Alignment Scores:
Pred. No.: 7,82e-34 Length: 371
Score: 376.00 Matches: 82
Percent Similarity: 56.67% Conservative: 37
Best Local Similarity: 39.05% Mismatches: 81
Query Match: 30.03% Indels: 10
DB: 13 Gaps: 7

US-09-830-144-1_COPY_408_1091 (1-684) x Q90ZY8 (1-371)

QY 4 GAGCTTCGGCAGTATCCCGTGAACCACTGATTAATATGTAAGCTTTATGAGCGTGC 63
Db 84 GluAlaGluIleLeuSerValLeuSerHisLysAsnIleIleGlnPheTyrGlyAlaIle 103

```

```

QY 64 TTGAATCCAGTG-----TGCTTCTGATGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db 104 LeuGluAlaProAsnAspGlyIleValThrGluTyrAlaSerArgGlySerLeuTyrGlu 123
QY 118 GTGCTGCATGTGTGCTGAACCACTTGCATVATTATATCTGCTGCCACCGCAATGATGGTGT 177
Db 124 TyrLeuSerSerAlaAspSer---GluGluMetAspMetAspGlnValMetThrTrpAla 142
QY 178 TTACAGTGTCCCAAGAGTGGCTTATCTTCACACATGCACACCAACCAAGCGCTAATTCAC 237
Db 143 MetGluIleAlaLysGlyMetHisTyrLeuHisAlaGluAlaProLeuLysValIleHis 162
QY 238 AGGACCTCGAACCACCACTTACTGCTGGTTGCGAGGGGACAGTCTCTAAAAATTTGT 297
Db 163 ArgAspLeuLysSerArgAsnValValLeuThrAla---AspAsnValLeuLysIleCys 181
QY 298 GATTTTGTGACAGCTGTGACATTGACACACATGCACCAATAACAAG-----GGAGT 351
Db 182 AspPheGly---AlaSerLysMetValSerHisThrThrHisMetSerLeuValGlyThr 200
QY 352 GCTGCTTGGATGCGACCTGAAGTTTGAAGGTAGTAAATTACAGTGAATAAATGTGACGTC 411
Db 201 PheProTirPheMetAlaProGluValIleGlnSerLeuProValSerGluThrCysAspThr 220
QY 412 TTCAGTGGGGTATTATCTTTGGGAAGTGAATAACGCGTCGGAAACCCCTTTGATGAGATT 471
Db 221 TyrSerTyrGlyValValLeuTirPgluMetLeuThrArgGluValProPheLys----- 238
QY 472 GGTGCGCCAGCTTCCGAATCANGTG---GCTGTTCATAATGTTACTGCACACCACTG 528
Db 239 GlyPheGluGlyLeuGlnValAlaTrpLeuValValGluLysHisGluArgProThrIle 258
QY 529 ATAAAAAATTTACCTAAGCCCATTCAGACCTGATGACTGCTGTTGGTCTAAAGATCCT 588
Db 259 ProSerSerCysProAlaSerPheAlaAspLeuMetArgCysTrpAsnAlaGluPro 278
QY 589 TCCAGCGCCCTTCAATGGAGGAAATTTGTG 618
Db 279 LysGluArgProGlnPheLysGlnIleLeu 288

RESULT 11
Q8MVR1 PRELIMINARY; PRT; 2631 AA.
AC Q8MVR1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cyclic GMP-binding protein C.
GN GBPC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA Bosgraaf L., Russcher H., Smith J.L., Wessels D., Solls D.R.,
RA Van Haastert P.J.M.;
RT "A novel cGMP signalling pathway mediating myosin phosphorylation and
RT chemotaxis in Dictyostelium.";
RL EMBO J. 0:0-0(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22008080; PubMed=12011437;
RA Goldberg J.M., Bosgraaf L., Van Haastert P.J., Smith J.L.;
RT "Identification of four candidate cGMP targets in Dictyostelium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:6749-6754(2002).
DR EMBL; AF481923; AAM34041.1; -.
DR InterPro; IPR000595; cNMP_binding.
DR InterPro; IPR000591; DEP.
DR InterPro; IPR004182; GRAM dom.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000651; RasGEFN.

```

DR	InterPro; IPR001895; RasGRF_CDC25.
DR	InterPro; IPR001806; Ras trnsfrmg.
DR	InterPro; IPR002290; Ser thr kinase.
DR	InterPro; IPR001245; Tyr_kinase.
DR	Pfam; PF00027; cNMP_binding; 2.
DR	Pfam; PF00610; DEP; 1.
DR	Pfam; PF02893; GRAM; 1.
DR	Pfam; PF00069; pkinase; 1.
DR	Pfam; PF00617; RasGEF; 1.
DR	Pfam; PF00618; RasGEPN; 1.
DR	PRINTS; PR00449; RASTRNSFRMNG.
DR	PRINTS; PR00109; TYRKINASE.
DR	Prodom; PD000001; Prot.kinase; 1.
DR	SMART; SM00049; DEP; 1.
DR	SMART; SM00568; GRAM; 1.
DR	SMART; SM00369; LRR_TYP; 2.
DR	SMART; SM00147; RasGEF; 1.
DR	SMART; SM00220; S_TKc; 1.
DR	SMART; SM00219; TyrcK; 1.
DR	PROSITE; PS50042; CNMP_BINDING_3; 2.
DR	PROSITE; PS00186; DRP; 1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW	ATP-binding; GTP-binding; Transferase.
SQ	SEQUENCE 2631 AA; 294124 MW; 5BB7AABD0A637701 CRC64;

Alignment Scores:

Pred. No.:	1.45e-33	Length:	2631
Score:	375.50	Matches:	88
Percent Similarity:	53.66%	Conservative:	44
Best Local Similarity:	35.77%	Mismatches:	75
Query Match:	29.9%	Indels:	39
DB:	5	Gaps:	9

```

US-09-830-144-1_COPY_408_1091 (1-684) x Q8MWV1 (1-2631)

QY      4   GAGCTTCGGCAGTTATCCCGTGTGAACCATCCTTAATTGTAAGCTTTATGGAGCCTGC 63
Db      947 GluValTyrMetSerGlyLeuAsnHisProAsnValMetLysIleSerGlyPheCys 966
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      64   TTGAATCCAGTGTCGTGTGTCATGGATATGCTCAAGGGGCTCTTATATAATGTGCTG 123
Db      967 IleclnProLeuCysMetAlaLeuGlutyrValargTyrGlySerLeutySerLeuLeu 986
          :||:|||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      124 CATGGTGCTGAACCATATGCCATATTACTGCTCCCCACGCAATGAGTGTGTTA--- 180
Db      987 Ser-----AsnSerSerIleGlulleSerTrpGlyLeuarg 998
          ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      181 -----CAGTGTCACAGGAGTGCTTATCTTACAGCATGCACACCCAAA 225
Db      999 LeuGlnIlealaSerGlulleAlaLysGlyMetGlnHisLeuHisSerHisAsnPro--- 1017
          :::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      226 GCGCTAAATCACAGGACCTGAAACACCACCAACTTACTGCTG-----GTTCAGGGGGG 279
Db      1018 ProValIleHisargaspLeuLysSerProAsnIleLeuleuAsnGlyIleThrGlugly 1037
          ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      280 -----ACAGTTCCTAAAAATTGTGATTTTGGTACAGCCTGTGCATTCACAGACACATG 333
Db      1038 GlnAsnSerValalathrIleleasppheGlyThrSer-----1050
          :|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      334 ACCAANTAACAGGGGAGTGTGCT-----TGCATGGCA 366
Db      1051 ThrAlaleutyryGlyGlyAlaAlaLeulleArgCysValAspGlnProLeuTrpLeugly 1070
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      367 CQTGAAGTTTTCAAGGTAGTAGTATACAGTGAANAATGTGAGCTTTCAGCTGGGTATT 426
Db      1071 ProgluValneuAlaglyThrAlatyrseryerGluProSerAspValtyrSerPheGlylle 1090
          |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY      427 ATTCTTTGGGAAGTGATACCGCTCGGAACCCCTTTGATGAG-----ATTGTGGCCCA 480
Db      1091 IleLeuTrpGlueutyrrThrArgAlahisProPheAspgluPheGlnPheGlnTrp 1110
          |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

```

Db 2263 TripleAlaProGluLeuLeuAsnLysThrLysGluAlaSerThrGlnThrAspValTyr 2282
QY 415 AGCTGGGGTATTATCTTTGGGAAGTATACAGCGTCGGAACCCCTTTGATGAGATTGGT 474
Db 2283 AlaPheGlyValIleLeuTyrGluLeuValThrArgLysAspTyrPheGlyGluIleGly 2302
QY 475 GGCCAGCTTCCGAATCATGATGGCTGTTCTATATGTTACTCGACCACTGATAAAA 534
Db 2303 PheMetThrLeu---IleGluGluLysValIleAsnGlyGluArgProLysIleProGlu 2321
QY 535 AATTACCTAAGCCCATTCAGAGCTGATGACTGCTGTTGTTGTTCTAAAGATCTCCAG 594
Db 2322 AspCysProGluMetTyrSerLysLeuIleValGluCysTyrGlnThrAspAlaSerGln 2341
QY 595 CGCCCTTCAATGGAGGAAATT 615
Db 2342 ArgProLysPheSerGluIle 2348
QY 595 CGCCCTTCAATGGAGGAAATT 615
Db 2342 ArgProLysPheSerGluIle 2348
RESULT 13
Q9ZQ31 PRELIMINARY; PRT; 411 AA.
AC Q9ZQ31: Q94A10;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
CX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Gibson H.A.,
RA Goldsmith A.D., Jiang P.X., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Yu G., Yu S., Bowser L., Carninci P., Chen H.,
RA Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RA "Full Length cDNA of gene T28124.9/At2g24360 (GI:4337195).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Yamada K., Chan H.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
RA Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
RA Hu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RL "Arabidopsis Open Reading Frame (ORF) Clones.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC EMBL; AC006403; AAD18109.2; -
DR EMBL; AY046026; AAK76700.1; -
DR EMBL; AY133876; AAM91810.1; -
DR HSSP; P12931; 1FMK.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr pkinese.
DR InterPro; IPR001245; Tyr pkinese.
DR Pfam; PF00069; pkinase; 1.

```

```

DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Hypothetical protein; ATP-binding, Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 411 AA; 46001 MW; 3B7001CFBB411BAA CRC64;

Alignment Scores:
Pred. No.: 1-36e-33 Length: 411
Score: 374.00 Matches: 86
Percent Similarity: 57.80% Conservative: 40
Best Local Similarity: 39.45% Mismatches: 70
Query Match: 29.87% Indels: 22
DB: 10 Gaps: 9

US-09-830-144-1_COPY_408_1091 (1-684) x Q9ZQ31 (1-411)
QY 4 GAGCTTCGGCAGTTATCCGCTGTGAACCATCTCTATATTTGTAAGCTTTTATGGAGCTGC 63
Db 179 GluValSerMetLeuAlaAsnLeuLysHisProAsnIleValArgPheIleGlyAlaCys 198
QY 64 TTGAATCCAGTC-----TGCTTTGTGATGAATATGCTGAAGGGGCTCTTTTATATAAT 117
Db 199 ArgLysProMetValTyrCysIleValThrGluTyrAlaLysGlySerValArgGln 218
QY 118 GTGCTG-----CATGCTGTGAACCATTCGCCATATTTATATCTGCTGCCACGCA 165
Db 219 PheLeuThrArgArgGlnAsnArgAlaValProLeuLysLeu-----Ala 233
QY 166 ATGAGTTGGTGTGTTACAGTGTTCGCCAAGGAGTGCTTATCTTCACAGCATGCACCCAAA 225
Db 234 ValLysGlnAlaLeuAspValAlaArgGlyMetAlaTyrValHisGly-----Arg 250
QY 226 GCCTTAATTCACAGGACCTGAACCCACCAACTTACTGCTGCTGTCAGGGGGACAGTT 285
Db 251 AsnPheIleHisArgAspLeuLysSerAspAsnLeuIleSerAlaAspLysSerIle 270
QY 286 CTAAAAATTTGTGATTTTGTGTACAGCC--TGTGACATTCAGACACAC---ATGACCAAT 339
Db 271 ---LysIleAlaAspPheGlyValAlaArgGluValGlnThrGluGlyMetThrPro 289
QY 340 AACAAAGGGAGTCTGCTGATGGCAGCTGACCTGAAGTTTTCAGAGTAGTATTACAGTAA 399
Db 290 GluThrGlyThrTyrArgTyrMetAlaProGluMetIleGlnHisArgAlaTyrAsnGln 309
QY 400 AATGTGACGCTCTTCAGCTGGGCTATATCTTTGGAGTGTATGACGCTCGGAAACCC 459
Db 310 LysValAspValTyrSerPheGlyIleValLeuTyrGluLeuIleThrGlyLeuLeuPro 329
QY 460 TTTGATGAGATTGCTGGC-----CCAGCTTTCCGATCATGTGGGCTGTTCATAATGGT 513
Db 330 PheGlnAsnMetThrAlaValGlnAlaAlaPheAlaVal-----ValAsnArgGly 346
QY 514 ACTCGACCACTGATATAAAAAATTTACCTAAGCCCATTCAGAGCTGTGAGCTGATCCTGT 573
Db 347 ValArgProThrValProAsnAspCysLeuProValLeuSerAspIleMetThrArgCys 366
QY 574 TGGTCTAAAGATCTTCCAGCGCCCTTCAATGGAGGAAATTTGAAAAAATG 627
Db 367 TrpAspAlaAsnProGluValArgProCysPheValGluValValLysLeuLeu 384

RESULT 14
Q9M085 PRELIMINARY; PRT; 412 AA.
ID Q9M085
AC Q9M085;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2003 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE protein kinase-like protein (AT4G31170/P6E21_90).
GN AT4G31170.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

US-09-830-144-1\_COPY\_408\_1091 (1-684) x Q9M085 (1-412)

QY 4 GAGCTTCGGCAGCTATCCGCTGTGAACCATCTTAATATTGTTAAAGCTTTATCGAGCGTGC 63  
||| ||| : : : : :  
Db 180 GluValSerMetLeuAlaPheLeuLysHisProAsnIleValArgPheIleGlyAlaCys 199  
||| ||| : : : : :  
QY 64 TTGAATCCAGTG-----TGCTTGTGTGAATAATGCTGAAGGGGGCTCTTTATATAAT 117  
||| ||| : : : : :  
Db 200 IleLysProMetValTipCysIleValThrGluTyAlaLysGlySerValArgGln 219  
||| ||| : : : : :  
QY 118 GTGCTG-----CATGGTCTGAACATTGCCAATTACTTACTGCTGCCACGCA 165  
||| ||| : : : ||| |||  
Db 220 PheLeuThrLysArgGlnAsnArgAlaValProLeuLysLeu-----Ala 234  
||| ||| : : : ||| |||  
QY 166 ATGAGTTGGTGTTCACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAA 225  
||| ||| : : : ||| ||| : : : ||| |||  
Db 235 ValMetGlnAlaLeuAspValAlaArgGlyMetAlaTyValHISglu-----Arg 251  
||| ||| : : : ||| ||| : : : ||| |||  
QY 226 GCGCTAATTCAGAGGAGCTCAAACCACCAAACTTACTGCTGGTTCAGAGGGGGACAGTT 285  
||| ||| : : : ||| ||| : : : ||| |||  
Db 252 AsnPheIleHisArgAspLeuLysSerAspAsnLeuLeuIleSerAlaAspArgSerIle 271  
||| ||| : : : ||| ||| : : : ||| |||  
QY 286 CTAAAAATTTGTGATTTGGFACAGCC---TGTGACATTTCAGACACAC--ATCACCAAT 339  
||| ||| : : : ||| ||| : : : ||| |||  
Db 272 ---LysIleAlaAspPheGlyValAlaAargileGluValGlnThrGluGlyMetThrPro 290  
||| ||| : : : ||| ||| : : : ||| |||  
QY 340 AACAAAGGGAGTGTCTGTGATGGCACCCTGAAGTTTTTGAAGTAGTAATTACAGTAA 399  
||| ||| : : : ||| ||| : : : ||| |||  
Db 291 GluThrGlyThyArgTrpMetAlaProGluMetIleGlnHisArgProTyThrGln 310  
||| ||| : : : ||| ||| : : : ||| |||  
QY 400 AAATGTGACGCTTCACGTGGGTATTATTCTTTGGGAAGTGTAAAGCGTCGGAACCCC 459  
||| ||| : : : ||| ||| : : : ||| |||  
Db 311 LysValaspValTySerPheGlyIleValLeuTrpGluLeuIleThrGlyLeuLeuPro 330  
||| ||| : : : ||| ||| : : : ||| |||  
QY 460 TTTGATGAGATCGGTGGC-----CCAGCTTTCGGAATCATGTGGGCTGTTCATAATGGT 513  
||| ||| : : : ||| ||| : : : ||| |||  
Db 331 PheGlnAsnMetThrAlaValGlnAlaPheAlaVal-----ValasnArgGly 347  
||| ||| : : : ||| ||| : : : ||| |||  
QY 514 ACTCGACCACCTGATMAAAATTTACTTAAGCCATTGAGAGCTGATGATCGTTGT 573  
||| ||| : : : ||| ||| : : : ||| |||  
Db 348 ValArgProThrValProAlaAspCysLeuProValLeuGlyLeuIleMetThrArgCys 367  
||| ||| : : : ||| ||| : : : ||| |||  
QY 574 TGCTCTAAAGATCCTTCCAGCGCCCTTCATCGAGGAAATTTGTGAAAATAATG 627  
||| ||| : : : ||| ||| : : : ||| |||  
Db 368 TrpAspAlaaspProGluValArgProCysPheAlaGluIleValAsnLeuLeu 385  
||| ||| : : : ||| ||| : : : ||| |||

RESULT 15

Q8GV30 PRELIMINARY; PERT; 417 AA.

ID Q8GV30 AC AGV30;  
DT 01-MAR-2003 (TEMBLrel. 23, Created)  
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)  
DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)  
GN Serine/threonine protein kinase.  
DE DPK1.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartioideae; Oryzaceae; Oryza.  
NCBI TaxID=39947;  
RN [1]\_TaxID=39947;  
RP SEQUENCE FROM N.A.  
RA Gu Z.M., Zhang H.S., Huang J.;  
RT "Molecular Cloning And Characterization Of A Novel Serine/Threonine  
RT Protein Kinase Gene Family From Rice.";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY156510; AAN84502.1; -.  
KW Kinase.  
SQ SEQUENCE 417 AA; 45927 MW; 1B66E3596653PAD7 CRC64;

Alignment Scores:  
Pred. No.: 1.77e-33 Length: 417  
Score: 373.00 Matches: 84

RT Molecular Cloning And Characterization Of A Novel Serine/Threonine  
 Kinase  
 RL Protein Kinase Gene Family From Rice."  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY156510; AAN84502.1; --  
 KW Kinase.  
 SQ SEQUENCE 417 AA; 45927 MW; 1B66E3596653FAD7 CRC64;  
 Alignment Scores:  
 Pred. No.: 1.77e-33 Length: 417  
 Score: 373.00 Matches: 84

Thu Dec 4 17:00:21 2003

Percent Similarity: 57.34% Conservative: 41  
 Best Local Similarity: 38.53% Mismatches: 71  
 Query Match: 29.79% Indels: 22  
 DB: 10 Gaps: 9

US-09-830-144-1\_COPY\_408\_1091 (1-684) x Q8GV30 (1-417)

```

QY 4 GAGCTTCGGCAGTTATCCGGTGTGACCATCTATATTTAAAGCTTTATGGAGCCTGC 63
Db 185 GluValMetLeuAlaThrLeuArgHisSerAsnIleValLysPheValGlyAlaCys 204
QY 64 TTGAATCCAGTG-----TGTCTTGTGATGAATATGCTGAAGGGGCTCTTTATATAAT 117
Db 205 ArgLysProMetValTrpCysIleValThrGluTyrAlaLysGlyGlySerValArgAsn 224
QY 118 GTGCTGCATGTGCTGAA-----CCATTGCCATATTATATCTCTGCCACGCA 165
Db 225 PheLeuAsnArgArgGlnAsnArgSerValProLeuLysLeu-----Ala 239
QY 166 ATGAGTTGTTTACAGTTGTTCCCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAA 225
Db 240 ValLysGlnAlaLeuAspValAlaArgGlyMetAlaTyrValHisGlyLeu----- 256
QY 226 GCGCTAATTACAGGACCTGAACACCAAACTTACTGTCTGTTCAGGGGGACAGTT 285
Db 257 GlyPheIleHisArgAspLeuLysSerAspAsn---LeuLeuIleSerGlyAspLysSer 275
QY 286 CTAAAAATTTGTCATTTGTACAGCC---TGTGACATTCAGACACAC---ATGACCAAT 339
Db 276 IleLysIleAlaAspPheGlyValAlaArgIleGluValLysThrGluGlyMetThrPro 295
QY 340 AACAGGGAGTCTGCTGATGGACCTGAGCTTTTGAAGGTAGTATTAACAGTGAA 399
Db 296 GluThrGlyThrTyrArgTrpMetAlaProGluValIleGlnHisArgProTyrAspGln 315
QY 400 AATGTGACCTCTTCAGCTGGGTATTATTCTTTGGGAAGTATAGCGCTCGGAAACCC 459
Db 316 LysValAspValTyrSerPheGlyIleValLeuTrpGluLeuValThrGlyAsnLeuPro 335
QY 460 TTTGATGAGATTGTGTGC-----CCAGCTTTCGAATCATGTGGGCTGTTCAATGGT 513
Db 336 PheAlaAsnMetThrAlaValGlnAlaAlaPheAlaVal-----ValAsnLysGly 352
QY 514 ACTGACACACCTGATAAATAATTTACCTAAGCCCATTTGAGAGCCTGATGACTCGTTGT 573
Db 353 ValArgProAlaIleProHisAspCysLeuProAlaLeuAlaGluIleMetThrArgCys 372
QY 574 TGGTCTAAAGATCCTTCCAGCGCCCTTCAATGGAGGAATTTGAAAAATAATG 627
Db 373 TrpAspAlaAsnProAspAlaArgProPheThrGluValValArgMetLeu 390

```

Search completed: December 4, 2003, 09:07:12  
 Job time : 86.7973 secs



1	684	100.0	1704	20	AA399698	Human TGF-beta act
2	684	100.0	1788	20	AA26285	Human TAK1-5XHis e
3	684	100.0	1959	18	AAT85095	Human transforming
4	684	100.0	2656	20	AA356279	Human TAK1 encodin
5	684	100.0	2656	21	AA339105	Human TAK-1 nucleo
6	684	100.0	2769	24	ABJ88437	Pain regulated cDN
7	684	100.0	2785	20	AA399696	Human TGF-beta act
8	684	100.0	2866	20	AA399697	Human TGF-beta act

PA	(TANA ) TANABE SEIYAKU CO.	
XX	Hasegawa K, Kageyama N, Sakurai H, Sugita T;	
PI	WPI: 1999-494298/41.	
XX	P-PSDB; AAY28998.	
DR	Nuclear factor kappa B activation inhibitors, useful as preventives	
PT	for, e.g. autoimmune diseases	
XX	Examples; Page 43-46; 49pp; Japanese.	
PS		
XX	The invention provides a method for identifying or screening a nuclear	
CC	factor kappa B (NF-kB) activation inhibitor by examining the effect of a	
CC	test substance on modulating the functions of TGF-beta activated kinase	
CC	1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to	
CC	treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),	
CC	intractable diseases with inflammation (such as atrophic dermatitis and	
CC	psoriasis), viral infection, endotoxin shock, septicemia and others. The	
CC	present sequence represents the nucleotide sequence of human TAK1C	
CC	(hTAK1C) protein.	
XX		
SQ	Sequence 1704 BP; 511 A; 381 C; 401 G; 411 T; 0 other;	
	Query Match 100.0%; Score 684; DB 20; Length 1704;	
	Best Local Similarity 100.0%; Pred. No. 2.3e-214; Indels 0; Gaps 0;	
	Matches 684; Conservative 0; Mismatches 0;	
QY	1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCACTCTTAATTTGTAAGCTTTTATGAGCC 60	
DB	226 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCACTCTTAATTTGTAAGCTTTTATGAGCC 285	
QY	61 TCGTTGAATCCAGTGTGCTTGTGATGGAATATGCTGGAAGGGGCTCTTTATATAATGTG 120	
DB	286 TCGTTGAATCCAGTGTGCTTGTGATGGAATATGCTGGAAGGGGCTCTTTATATAATGTG 345	
QY	121 CTGCATGGTCTGAACATGTCATATATATCTGCTGCCACGCAATGATGGTGTGTTA 180	
DB	346 CTGCATGGTCTGAACATGTCATATATATCTGCTGCCACGCAATGATGGTGTGTTA 405	
QY	181 CAGTGTTCACAGAGTGGCTTATCTTCACAGATGCAACCCAAAGCGCTTAATTCACAGG 240	
DB	406 CAGTGTTCACAGAGTGGCTTATCTTCACAGATGCAACCCAAAGCGCTTAATTCACAGG 465	
QY	241 GACCTGAACACCAAACTTACTGCTGGTTCAGGGGGAGAGTCTTAAATAATGTGAT 300	
DB	466 GACCTGAACACCAAACTTACTGCTGGTTCAGGGGGAGAGTCTTAAATAATGTGAT 525	
QY	301 TTTGTACAGCCTGTGACATTCACACATGACCAATACCAAGGGAGTGGTCTTGG 360	
DB	526 TTTGGTACAGCCTGTGACATTCACACATGACCAATACCAAGGGAGTGGTCTTGG 585	
QY	361 ATGGCCACCTGAAGTTTGTGAAGTAGTAATTAACAGTGAAAAATGTGACGTCTTCAGCTGG 420	
DB	586 ATGGCCACCTGAAGTTTGTGAAGTAGTAATTAACAGTGAAAAATGTGACGTCTTCAGCTGG 645	
QY	421 GGTATATTTCTTGGGAAGTGAATACCGCTCGGAACCCCTTGTATGATGGTGGCCCA 480	
DB	646 GGTATATTTCTTGGGAAGTGAATACCGCTCGGAACCCCTTGTATGATGGTGGCCCA 705	
QY	481 GCTTTCGAATCATGTGGCTGTTCAATAATGTTACTGCACACCACTGATAAAAAATTTA 540	
DB	706 GCTTTCGAATCATGTGGCTGTTCAATAATGTTACTGCACACCACTGATAAAAAATTTA 765	
QY	541 CCTAAGCCCATTTGAGAGCCTGATGACCTGTTGTTGGTCTAAAGATCCTTCCAGCGCCCT 600	
DB	766 CCTAAGCCCATTTGAGAGCCTGATGACCTGTTGTTGGTCTAAAGATCCTTCCAGCGCCCT 825	
QY	601 TCAATGAGGAATTTGTGAATAATGACTCCTGATGGGTACTTTCAGAGGACAGAT 660	
DB	826 TCAATGAGGAATTTGTGAATAATGACTCCTGATGGGTACTTTCAGAGGACAGAT 885	
QY	561 GAGCCATTACAGTATCTTGTGCAG 684	

DB

886 GAGCCATTACAGTATCTTGTGCAG 909

RESULT 2

AA56285

ID AAX56285 standard, DNA; 1788 BP.

XX AC AAX56285;

XX 21-JUL-1999 (first entry)

XX Human TAK1-6xHis encoding DNA.

XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;

XX transforming growth factor beta; ss.

XX Homo sapiens.

XX Synthetic.

XX Key Location/Qualifiers

XX CDS 7..1779

XX /\*tag= a

XX WO9921010-A1.

XX 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS ) CHUGAI SEIYAKU KK.

XX Ohtomo T, Ono K, Tsuchiya M;

XX WPI: 1999-312645/26.

XX P-PSDB; AAY09547.

XX Screening for TGF- beta inhibitory substances, which are useful as

XX drugs for treatment of diseases relating to its disorder

XX Example 1; Page 167-171; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit

XX the binding of TAK1 polypeptide to TAB1 polypeptide. The method

XX comprises: (a) contacting the polypeptide in the presence of a sample;

XX and (b) detecting the amount of bound polypeptide, in which the sample

XX can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming

XX growth factor (TGF)-beta inhibitory substances can be used in drugs for

XX indications e.g. as TGF-beta signal transduction inhibitors or

XX activators, or extracellular matrix protein production enhancement

XX inhibitors or activators, or cell proliferation prevention inhibitors or

XX activators, or monocyte migration inhibitors or activators, or

XX physiological activity induction inhibitors or activators, or

XX immunosuppression inhibitors or activators, and such substances can also be

XX precipitation inhibitors or activators, and such substances can also be

XX inhibitors of the TAK1 polypeptide function, particularly kinase

XX activity. The present sequence encodes TAK1-6xHis from an example of

XX the present invention.

XX Sequence 1788 BP; 551 A; 394 C; 413 G; 430 T; 0 other;

XX

Query Match

100.0%; Score 684; DB 20; Length 1788;

Best Local Similarity 100.0%; Pred. No. 2.3e-214;

Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCACTCTTAATTTGTAAGCTTTTATGAGCC 60

DB

232 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCACTCTTAATTTGTAAGCTTTTATGAGCC 291

QY

61 TCGTTGAATCCAGTGTGCTTGTGATGGAATATGCTGGAAGGGGCTCTTTATATAATGTG 120

Db 292 TGCTTGATCCAGTGTGTCTGTGATGATATGCTGAAGGGGCTCTTTATATAATGTG 351  
QY 121 CTGATGTGTGTGAACCAATTCGCATATATATCTGCTGCCACGCAATGAGTTGGTGTTA 180  
Db 352 CTGATGTGTGTGAACCAATTCGCATATATATCTGCTGCCACGCAATGAGTTGGTGTTA 411  
QY 181 CAGTGTTCCTCAAGAGTGGCTTATCTTCCACAGCATGCAACCAAGCGCTAATTCACAGG 240  
Db 412 CAGTGTTCCTCAAGAGTGGCTTATCTTCCACAGCATGCAACCAAGCGCTAATTCACAGG 471  
QY 241 GACCTGAACCAACCAACTTACTGCTGTGTCAGGGGGAAGTCTTCAAAAATTTGTGAT 300  
Db 472 GACCTGAACCAACCAACTTACTGCTGTGTCAGGGGGAAGTCTTCAAAAATTTGTGAT 531  
QY 301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATACAGGGGAGTGTCTTGG 360  
Db 532 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATACAGGGGAGTGTCTTGG 591  
QY 361 ATGGCACTGAAGTTTGTGAAGGTAGTAATACAGTGAAGAAATGTGACGCTTCCAGCTGG 420  
Db 592 ATGGCACTGAAGTTTGTGAAGGTAGTAATACAGTGAAGAAATGTGACGCTTCCAGCTGG 651  
QY 421 GGTATTATCTTTGGAGTGAATACCGGTGCGAAACCTTTGTATGAGATTGGTGCCCA 480  
Db 652 GGTATTATCTTTGGAGTGAATACCGGTGCGAAACCTTTGTATGAGATTGGTGCCCA 711  
QY 481 GCTTTCCGAATCATGTGGCTGTTCATAATGTGATCTGACACACACATGATATAAAAAATTA 540  
Db 712 GCTTTCCGAATCATGTGGCTGTTCATAATGTGATCTGACACACACATGATATAAAAAATTA 771  
QY 541 CCTAAGCCCATTTGAGCCCTGATGACTCTGTTTGGTCTAAGATCCTTCCAGCGCCT 600  
Db 772 CCTAAGCCCATTTGAGCCCTGATGACTCTGTTTGGTCTAAGATCCTTCCAGCGCCT 831  
QY 601 TCAATGAGAAATTTGAAATAATGACTCACTTGATGCGGTACTTCCAGGAGCAGAT 660  
Db 832 TCAATGAGAAATTTGAAATAATGACTCACTTGATGCGGTACTTCCAGGAGCAGAT 891  
QY 661 GAGCCATTACAGTATCCTTGTGAG 684  
Db 892 GAGCCATTACAGTATCCTTGTGAG 915

## RESULT 3

AA8785095

ID AA8785095 standard; cDNA; 1959 BP.

XX AC AA8785095;

XX DT 19-NOV-1997 (first entry)

XX DE Human transforming growth factor-beta activated kinase TAK-1 cDNA.

XX DX TGF-beta; signal transduction; TGF-beta activated kinase;

XX EX MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;

XX KW protein kinase; ss.

XX OS Homo sapiens.

XX FH Key

XX FT CDS

XX FT 183..1922

XX FT /\*tag= a

XX FT /product= TAK-1

XX PN Location/Qualifiers

XX PD 24-JUN-1997.

XX PF 27-SEP-1996; 96JP-0256747.

XX PR 24-JUL-1996; 96US-0685625.

XX PR 29-SEP-1995; 95JP-0253549.

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

PA (CHUS) CHUGAI PHARM CO LTD.

PA (UENO) UENO N.

XX WPI; 1997-380171/35.

XX P-PSDB; AAW27093.

XX DNA encoding transforming growth factor-beta-activated kinase, TAK-1

XX - useful for studying the TGF-beta signal transduction system

XX Claim 9; Page 13-15; 20pp; Japanese.

XX The present sequence encodes human transforming growth factor-beta

XX (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the

XX TAK-1 protein which is involved in the TGF-beta family signal

XX transduction system. TAK-1, also known as activator of MAPK Kinase

XX (AMK-1), is an enzyme which is activated by TGF-beta and bone

XX morphogenetic protein (BMP) and activates MAPK kinase by

XX phosphorylation.

XX Sequence 1959 BP; 571 A; 447 C; 481 G; 460 T; 0 other;

XX Query Match 100.0%; Score 684; DB 18; Length 1959;

XX Best Local Similarity 100.0%; Pred. No. 2.4e-214;

XX Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCACTCTAATATGTGTAAGCTTTATGAGCC 60

Db 408 GTAGAGCTTCGGCAGTATCCCGTGTGAACCACTCTAATATGTGTAAGCTTTATGAGCC 467

QY 61 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120

Db 468 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 527

QY 121 CTGATGTGTCTGAACCACTTGCATATATATCTGCTCCACGCAATGAGTTGGTGTTA 180

Db 528 CTGATGTGTCTGAACCACTTGCATATATATCTGCTCCACGCAATGAGTTGGTGTTA 587

QY 181 CAGTGTTCCTCAAGAGTGGCTTATCTTCAAGCATGCAACCAAGCGCTTAATTCACAGG 240

Db 588 CAGTGTTCCTCAAGAGTGGCTTATCTTCAAGCATGCAACCAAGCGCTTAATTCACAGG 647

QY 241 GACCTGAACCAACCAACTTACTGCTGTGAGGGGGGACAGTTCTTAAAAATTTGTGAT 300

Db 648 GACCTGAACCAACCAACTTACTGCTGTGAGGGGGGACAGTTCTTAAAAATTTGTGAT 707

QY 301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATAACAGGGGAGTGTCTTGG 360

Db 708 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATAACAGGGGAGTGTCTTGG 767

QY 361 ATGGCACTGAAGTTTGTGAAGGTAGTAATACAGTGAAGAAATGTGACGCTTTTCAAGCTGG 420

Db 768 ATGGCACTGAAGTTTGTGAAGGTAGTAATACAGTGAAGAAATGTGACGCTTTTCAAGCTGG 827

QY 421 GGTATTATCTTTGGAGTGAATACCGCTGCGAAACCTTTGATGAGATTGGTGCCCA 480

Db 828 GGTATTATCTTTGGAGTGAATACCGCTGCGAAACCTTTGATGAGATTGGTGCCCA 887

QY 481 GCTTTCCGAATCATGTGGCTGTTCATAATGTGATCTGACACACACATGATATAAAAAATTA 540

Db 888 GCTTTCCGAATCATGTGGCTGTTCATAATGTGATCTGACACACACATGATATAAAAAATTA 947

QY 541 CCTAAGCCCATTTGAGCCCTGATGACTCTGTTTGGTCTTAAAGATCCTTCCAGCGCCT 600

Db 948 CCTAAGCCCATTTGAGCCCTGATGACTCTGTTTGGTCTTAAAGATCCTTCCAGCGCCT 1007

QY 601 TCAATGAGAAATTTGAAATAATGACTCACTTGATGCGGTACTTCCAGGAGCAGAT 660

Db 1008 TCAATGAGAAATTTGAAATAATGACTCACTTGATGCGGTACTTCCAGGAGCAGAT 1067

QY 661 GAGCCATTACAGTATCCTTGTGAG 684

Db 1068 GAGCCATTACAGTATCCTTGTGAG 1091

```

RESULT 4
AAX56279
ID AAX56279 standard; DNA; 2656 BP.
XX
AC AAX56279;
XX
DT 21-JUL-1999 (first entry)
XX
DE Human TAK1 encoding DNA.
XX
KW Human; TAK1; screening; inhibition; TGF-beta;
transforming growth factor beta; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 183..1922
FT /*tag= a
XX
PN WO9921010-A1.
XX
PD 29-APR-1999.
XX
PF 22-OCT-1998; 98WO-JP04796.
XX
PR 22-OCT-1997; 97JP-0290188.
XX
PA (CHUS ) CHUGAI SEIYAKU KK.
XX
PI Ohtomo T, Ono K, Tsuchiya M;
XX
DR WPI; 1999-312645/26.
XX
DR P-PSDB; AAY09542.
XX
PT Screening for TGF- beta inhibitory substances, which are useful as
drugs for treatment of diseases relating to its disorder
XX
PS Example 1; Page 150-154; 195pp; Japanese.
XX
CC A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes human TAK1.
XX
SQ Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;
Query Match 100.0%; Score 684; DB 20; Length 2656;
Best Local Similarity 100.0%; Pred. No. 2.8e-214;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCACTTAATTTGTAAGCTTTATGGAGCC 60
Db 408 GTAGAGCTTCGGCAGTATCCCGTGTGAACCACTTAATTTGTAAGCTTTATGGAGCC 467
QY 61 TCGTTGAATCCAGTGTCTTGTGTGATGGAATATGCTCAAGGGGCTCTTTATATAATGTG 120
Db 468 TCGTTGAATCCAGTGTCTTGTGTGATGGAATATGCTCAAGGGGCTCTTTATATAATGTG 527
QY 121 CTGCATGCTGTGAACCAATTGCCATATTACTCTGCCACGCAATGAGTGTGTTTA 180
Db 528 CTGCATGCTGTGAACCAATTGCCATATTACTCTGCCACGCAATGAGTGTGTTTA 587

```

```

181 CAGTGTTCCTCCAGAGTGGCTTATCTTCACAGCATGCAACCAAGCCGCTAATTCACAGG 240
588 CAGTGTTCCTCCAGAGTGGCTTATCTTCACAGCATGCAACCAAGCCGCTAATTCACAGG 647
241 GACCTGAAACCAACCAAACTTACTGCTGTGTCAGGGGGACAGTCTTAAAAATTTGTGAT 300
648 GACCTGAAACCAACCAAACTTACTGCTGTGTCAGGGGGACAGTCTTAAAAATTTGTGAT 707
301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATTAACAGGGGAGTCTGCTTGG 360
708 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATTAACAGGGGAGTCTGCTTGG 767
361 ATGGCACCTGAAATTTTGAAGGTAGTAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 420
768 ATGGCACCTGAAATTTTGAAGGTAGTAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 827
421 GGTATTATCTTTGGGAAGTGTAAACGGTCCGAAACCTTTGATGAGATTGGTGCCCA 480
828 GGTATTATCTTTGGGAAGTGTAAACGGTCCGAAACCTTTGATGAGATTGGTGCCCA 887
481 GCTTTCCGAATCATGTGGGCTCTTCATATGCTTACTCGACCACTGATATAAAAAATTTA 540
889 GCTTTCCGAATCATGTGGGCTCTTCATATGCTTACTCGACCACTGATATAAAAAATTTA 947
541 CCTAAGCCCATTCAGAGCCTGATGACTCGTTGTTGTTCTAAAGATCCTTCCAGCGCCCT 600
948 CCTAAGCCCATTCAGAGCCTGATGACTCGTTGTTGTTCTAAAGATCCTTCCAGCGCCCT 1007
601 TCAATGGAGGAAATTTGTAATAATGACTCTGATCGGGTACTTCCAGGAGCAGAT 660
1008 TCAATGGAGGAAATTTGTAATAATGACTCTGATCGGGTACTTCCAGGAGCAGAT 1067
661 GAGCCATTACAGTATCTTGTCTGAG 684
1068 GAGCCATTACAGTATCTTGTCTGAG 1091

```

## RESULT 5

AAX39105

ID AAA39105 standard; DNA; 2656 BP.

XX

AC AAA39105;

XX

DT 04-SEP-2000 (first entry)

XX

DE Human TAK-1 nucleotide sequence SEQ ID NO:1.

XX

KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;

KW screening; signal transduction; inhibition; inflammatory cytokine;

KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;

KW antiinflammatory; suppression; ds.

XX

XX Homo sapiens.

XX

FH Key Location/Qualifiers

FT 183..1922

FT /\*tag= a

FT /product= "TAK-1"

XX

XX WO200023610-A1.

XX

XX 27-APR-2000.

XX

XX 21-OCT-1999; 99WO-JP05817.

XX

XX 21-OCT-1998; 98JP-0299962.

XX

XX (CHUS ) CHUGAI SEIYAKU KK.

XX

XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;

XX

XX WPI; 2000-339707/29.

XX

XX

XX

XX

XX

XX

XX

DR P-PSDB; AAY91000.  
 XX Method for screening inhibitors of TAK1 signal transduction for  
 PT suppression of inflammatory cytokine production and use as  
 PT antiinflammatory agents  
 XX  
 PS Example 1; Page 73-80; 100pp; Japanese.  
 XX  
 CC The present invention describes a method for screening compounds for  
 CC inhibition of inflammatory cytokine signal transduction by contacting  
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition  
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds  
 CC for inhibition of inflammatory cytokine signal transduction in which the  
 CC inhibition of TAK1 phosphorylation is selected for; and drug  
 CC compositions for the treatment of inflammatory disorders containing as  
 CC active component an inflammatory cytokine signal transduction inhibitor.  
 CC TAK1 is an essential component of the signalling process which results  
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),  
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used  
 CC for the selection of effective antiinflammatory agents. The present  
 CC sequence encodes human TAK-1, which is used in the exemplification of the  
 CC present invention.  
 XX  
 SQ Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;  
 Query Match 100.0%; Score 684; DB 21; Length 2656;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-214;  
 Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 60  
 408 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 467  
 61 TCGTTGAATCCAGTGTCTGTGATCGAATATGCTGAAGGGGCTCTTTATAAATGTG 120  
 468 TCGTTGAATCCAGTGTCTGTGATCGAATATGCTGAAGGGGCTCTTTATAAATGTG 527  
 121 CTGCATGGTGTGAACCATTCGCCATATTATATCTGCTGCCACGCAATGAGTGTGTTA 180  
 528 CTGCATGGTGTGAACCATTCGCCATATTATCTGCTGCCACGCAATGAGTGTGTTA 587  
 181 CAGTGTTCGCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGG 240  
 588 CAGTGTTCGCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTTAATTCACAGG 647  
 241 GACCTGAACACCAACACTTACTGCTGGTTCAGGGGGGACAGTCTAAATAATTTGTGAT 300  
 648 GACCTGAACACCAACACTTACTGCTGGTTCAGGGGGGACAGTCTAAATAATTTGTGAT 707  
 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATACCAAGGGGAGTGGCTTGG 360  
 708 TTTGGTACAGCTGTGACATTCAGACACATGACCAATACCAAGGGGAGTGGCTTGG 767  
 361 ATGGCACTGAAGTTTGAAGTAGTAAATTCAGTGAATAATGTGACGCTTTCAGCTGG 420  
 768 ATGGCACTGAAGTTTGAAGTAGTAAATTCAGTGAATAATGTGACGCTTTCAGCTGG 827  
 421 GGTATTATCTTTGGGAGTAAACCGCTGCGAAACCTTTGATGAGATTGGTGGCCCA 480  
 828 GGTATTATCTTTGGGAGTAAACCGCTGCGAAACCTTTGATGAGATTGGTGGCCCA 887  
 481 GCTTCCGGAATCATGTGGGCTGTTTATAATGTTACTGACACCACTGATAAATAATTTA 540  
 888 GCTTCCGGAATCATGTGGGCTGTTTATAATGTTACTGACACCACTGATAAATAATTTA 947  
 541 CCTAAGCCCATTTGAGAGCTGATGATCGTGTGTGTCTAAAGATCCTTCCAGCGCCCT 600  
 948 CCTAAGCCCATTTGAGAGCTGATGATCGTGTGTGTCTAAAGATCCTTCCAGCGCCCT 1007  
 601 TCAATGAGGAATCTGAAATAATGACTCACTTGTATCGGGTATCTTCAGGAGCAGAT 660  
 1008 TCAATGAGGAATCTGAAATAATGACTCACTTGTATCGGGTATCTTCAGGAGCAGAT 1067

QY 661 GAGCCATTACAGTATCCTTGTGTCAG 684  
 DB 1068 GAGCCATTACAGTATCCTTGTGTCAG 1091  
 RESULT 6  
 ABL88437  
 ID ABL88437 standard; cDNA; 2769 BP.  
 AC ABL88437;  
 XX  
 DT 16-MAY-2002 (first entry)  
 XX  
 DE Pain regulated cDNA sequence 80.  
 XX  
 KW Pain; analgesic; gene therapy; neurological disorder;  
 KW neurodegenerative disease; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200212338-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 03-AUG-2001; 2001WO-EP09011.  
 XX  
 PR 03-AUG-2000; 2000DE-1037759.  
 XX  
 PA (CHEF ) GRUENENTHAL GMBH.  
 XX  
 PI Gillen C, Wetzels I, Wnendt S, Weihe E, Schaefer MK;  
 XX  
 DR WPI; 2002-257469/30.  
 DR P-PSDB; ABB85033.  
 XX  
 PT Identifying pain-regulating compounds, useful for treating chronic pain  
 PT and for diagnosis, by measuring binding of compounds to specific  
 PT peptides and proteins  
 XX  
 PS Claim 1; Fig 44; 213pp; German.  
 XX  
 CC The invention relates to identifying pain-regulating substances (A)  
 CC comprises (i) incubating a test substance with a cell (or preparation  
 CC from it) that has synthesised a peptide or protein (B) and (ii) measuring  
 CC either binding of the test substance to (B) or some functional parameter  
 CC that is altered by this binding. The method is useful for identifying  
 CC pain-regulating substances (A) with analgesic activity. (A) along with  
 CC nucleic acid (ABL88411-ABL88441) that encode proteins (B,  
 CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the  
 CC nucleic acid; antibodies against (B); cells that express (B) and agents  
 CC that bind to (B), are all useful for treating pain, particularly chronic  
 CC pain, including use in gene therapy. The same materials can also be used  
 CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The  
 CC present sequence is that of a polynucleotide of the invention.  
 XX  
 SQ Sequence 2769 BP; 811 A; 565 C; 640 G; 753 T; 0 other;  
 Query Match 100.0%; Score 684; DB 24; Length 2769;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-214;  
 Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 60  
 388 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 447  
 61 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATAAATGTG 120  
 448 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATAAATGTG 507  
 121 CTGCATGGTGTGAACCATTCGCCATATTATCTGCTGCCACGCAATGAGTGTGTTA 180  
 508 CTGCATGGTGTGAACCATTCGCCATATTATCTGCTGCCACGCAATGAGTGTGTTA 567

181 CAGTGTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 240  
 568 CAGTGTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCACAGG 627  
 241 GACCTGGAACCAACCAACTTACTGCTGTTGAGGGGGGACAGTCTTAAAAATTTGTGAT 300  
 628 GACCTGGAACCAACCAACTTACTGCTGTTGAGGGGGGACAGTCTTAAAAATTTGTGAT 687  
 301 TTGTGTACAGCTGTGACATTCAGACACACATGACCAATTAACAAGGGAGTCTGCTGG 360  
 688 TTGTGTACAGCTGTGACATTCAGACACACATGACCAATTAACAAGGGAGTCTGCTGG 747  
 361 ATGGCAGCTGAAGTTTTCAGAGTACTTAATTCAGTGAAGAAATGACGCTCTTCAGCTGG 420  
 748 ATGGCAGCTGAAGTTTTCAGAGTACTTAATTCAGTGAAGAAATGACGCTCTTCAGCTGG 807  
 421 GGTATTAATTTCTTGGGAAGTATGATACGCTGCGGAAACCTTTGATGAGATTGGTGGCCCA 480  
 808 GGTATTAATTTCTTGGGAAGTATGATACGCTGCGGAAACCTTTGATGAGATTGGTGGCCCA 867  
 481 GCTTTCCGAATCATGTGGGCTGTTCATATGATGTTGATGAGATTGGTGGCCCA 540  
 868 GCTTTCCGAATCATGTGGGCTGTTCATATGATGTTGATGAGATTGGTGGCCCA 927  
 541 CCTAAGCCCAATGAGAGCCTGTGATGCTGCTGTTGCTTAAAGATCCTTCCAGCGCCCT 600  
 928 CCTAAGCCCAATGAGAGCCTGTGATGCTGCTGTTGCTTAAAGATCCTTCCAGCGCCCT 987  
 601 TCAATGAGGAAATTTGTAAGAAATTAATGACTCACTTGAATGCGGTACTTTCCAGGAGCAGAT 660  
 988 TCAATGAGGAAATTTGTAAGAAATTAATGACTCACTTGAATGCGGTACTTTCCAGGAGCAGAT 1047  
 661 GAGCCATTACAGTATCCTTGTGAC 684  
 1048 GAGCCATTACAGTATCCTTGTGAC 1071

RESULT 7  
 AAX99696  
 ID AAX99696 standard; cdna to mRNA; 2785 BP.  
 AC AAX99696;  
 XX  
 DT 18-OCT-1999 (first entry)  
 XX  
 DE Human TGF-beta activated kinase (TAK) 1a encoding nucleotide sequence.  
 XX  
 KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;  
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;  
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;  
 KW endotoxin shock; septicemia; human; hTAK1a; ss.  
 XX  
 OS Homo sapiens.  
 PH Key Location/Qualifiers  
 FT CDS 163..1902  
 FT /\*tag= a  
 FT /\*product= "hTAK1a"  
 XX  
 WO9940202-A1.  
 XX  
 PD 12-AUG-1999.  
 XX  
 PF 02-FEB-1999; 99WO-JP00422.  
 XX  
 PR 30-OCT-1998; 98JP-0309316.  
 PR 06-FEB-1998; 98JP-0026003.  
 XX  
 PA (TANA ) TANABE SEIYAKU CO.  
 XX  
 XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;  
 XX WPI, 1999-494298/41.  
 DR

DR P-PSDB; AAY28996.  
 XX  
 PT Nuclear factor kappa B activation inhibitors, useful as preventives  
 for, e.g. autoimmune diseases  
 XX  
 PS Examples; Page 35-39; 49pp; Japanese.  
 CC  
 CC The invention provides a method for identifying or screening a nuclear  
 factor kappa B (NF-kB) activation inhibitor by examining the effect of a  
 test substance on modulating the function(s) of TGF-beta activated kinase  
 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
 treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
 CC intractable diseases with inflammation (such as atrophic dermatitis and  
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The  
 CC present sequence represents the nucleotide sequence of human TAK1a  
 CC (hTAK1a) protein.  
 XX  
 XX Sequence 2785 BP; 827 A; 565 C; 640 G; 753 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 684; DB 20; Length 2785;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-214;  
 Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCCTAATATTGTAAGCTTTATGAGCC 60  
 DB 388 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCCTAATATTGTAAGCTTTATGAGCC 447  
 QY 61 TGCTTGAATCCAGTGTCTGTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 120  
 DB 448 TGCTTGAATCCAGTGTCTGTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 507  
 QY 121 CTGCAATGCTGAACCATTCGCAATTAATGCTGTGCGCCACGCAATGAGTGTGTTA 180  
 DB 508 CTGCAATGCTGAACCATTCGCAATTAATGCTGTGCGCCACGCAATGAGTGTGTTA 567  
 QY 181 CAGTGTTCAGAGAGTGTCTTATCTTCAGCATGCAACCAAGCGCTTAATTCACAGG 240  
 DB 568 CAGTGTTCAGAGAGTGTCTTATCTTCAGCATGCAACCAAGCGCTTAATTCACAGG 627  
 QY 241 GACCTGAACCAACCAACTTACTGCTGTTGAGGGGGGACAGTCTTAAAAATTTGTGAT 300  
 DB 628 GACCTGAACCAACCAACTTACTGCTGTTGAGGGGGGACAGTCTTAAAAATTTGTGAT 687  
 QY 301 TTGTGTACAGCTGTGACATTCAGACACACATGACCAATTAACAAGGGAGTGTGCTGG 360  
 DB 688 TTGTGTACAGCTGTGACATTCAGACACACATGACCAATTAACAAGGGAGTGTGCTGG 747  
 QY 361 ATGGCAGCTGAAGTTTTCAGAGTACTTAATGATGTTGATGAGATTGGTGGCCCA 420  
 DB 748 ATGGCAGCTGAAGTTTTCAGAGTACTTAATGATGTTGATGAGATTGGTGGCCCA 807  
 QY 421 GGTATTAATTTCTTGGGAAGTATGATACGCTGCGGAAACCTTTGATGAGATTGGTGGCCCA 480  
 DB 808 GGTATTAATTTCTTGGGAAGTATGATACGCTGCGGAAACCTTTGATGAGATTGGTGGCCCA 867  
 QY 481 GCTTTCCGAATCATGTGGGCTGTTCATATGATGTTGATGAGATTGGTGGCCCA 540  
 DB 868 GCTTTCCGAATCATGTGGGCTGTTCATATGATGTTGATGAGATTGGTGGCCCA 927  
 QY 541 CCTAAGCCCAATGAGAGCCTGTGATGCTGCTGTTGCTTAAAGATCCTTCCAGCGCCCT 600  
 DB 928 CCTAAGCCCAATGAGAGCCTGTGATGCTGCTGTTGCTTAAAGATCCTTCCAGCGCCCT 987  
 QY 601 TCAATGAGGAAATTTGTAAGAAATTAATGACTCACTTGAATGCGGTACTTTCCAGGAGCAGAT 660  
 DB 988 TCAATGAGGAAATTTGTAAGAAATTAATGACTCACTTGAATGCGGTACTTTCCAGGAGCAGAT 1047  
 QY 661 GAGCCATTACAGTATCCTTGTGAC 684  
 DB 1048 GAGCCATTACAGTATCCTTGTGAC 1071

RESULT 8

AAX99697  
ID AAX99697 standard; cDNA to mRNA; 2866 BP.  
AC AAX99697;  
XX  
DT 18-OCT-1999 (first entry)  
XX  
DE Human TGF-beta activated kinase (TAK) 1b encoding nucleotide sequence.  
XX  
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;  
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;  
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;  
KW endotoxin shock; septicemia; human; hTAK1b; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 163..1983  
FT /\*tag= a  
FT /product= "hTAK1b"  
XX  
XX WO9940202-A1.  
XX  
PD 12-AUG-1999.  
XX  
PF 02-FEB-1999; 99WO-JP00422.  
XX  
PR 30-OCT-1998; 98JP-0309316.  
PR 06-FEB-1998; 98JP-0026003.  
XX  
XX (TANA ) TANABE SEIYAKU CO.  
XX  
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;  
XX WPI; 1999-494298/41.  
XX P-PSDB; AAY28997.  
XX  
XX Nuclear factor kappa B activation inhibitors, useful as preventives  
XX for, e.g. autoimmune diseases  
XX  
XX Examples; Page 39-43; 49pp; Japanese.  
XX  
XX The invention provides a method for identifying or screening a nuclear  
XX factor kappa B (NF-kB) activation inhibitor by examining the effect of a  
XX test substance on modulating the function(s) of TGF-beta activated kinase  
XX 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to  
XX treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),  
XX intractable diseases with inflammation (such as atrophic dermatitis and  
XX psoriasis), viral infection, endotoxin shock, septicemia and others. The  
XX present sequence represents the nucleotide sequence of human TAK1b  
XX (hTAK1b) protein.  
XX  
XX Sequence 2866 BP; 848 A; 588 C; 656 G; 774 T; 0 other;  
XX  
XX Query Match 100.0%; Score 684; DB 20; Length 2866;  
XX Best Local Similarity 100.0%; Pred. No. 2.9e-214;  
XX Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATTTAAAGCTTTATGAGCC 60  
DB 388 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATTTAAAGCTTTATGAGCC 447  
QY 61 TCGTTGAATCCAGTGTCTTTGTGAGGAATATCTGCTGCCCGCCCAATGAGTTGGTGTTA 180  
DB 448 TCGTTGAATCCAGTGTCTTTGTGAGGAATATCTGCTGCCCGCCCAATGAGTTGGTGTTA 120  
QY 121 CTGCATGGTCTGAACCATTTGCCATATTATCTGCTGCCCGCCCAATGAGTTGGTGTTA 507  
DB 508 CTGCATGGTCTGAACCATTTGCCATATTATCTGCTGCCCGCCCAATGAGTTGGTGTTA 567  
QY 181 CAGTGTTCACAGGAGTGGCTTATCTTCACAGCATGCCAACCCAAAGCGCTAATTCACAGG 240  
DB 568 CAGTGTTCACAGGAGTGGCTTATCTTCACAGCATGCCAACCCAAAGCGCTAATTCACAGG 627

QY 241 GACCTGAACACCAAACTTACTGCTGGTTCAGGGGGACAGTTCTAAAAATTTTGAT 300  
DB 628 GACCTGAACACCAAACTTACTGCTGGTTCAGGGGGACAGTTCTAAAAATTTTGAT 587  
QY 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATAACAGGGGAGTCTGCTTGG 360  
DB 688 TTTGGTACAGCTGTGACATTCAGACACATGACCAATAACAGGGGAGTCTGCTTGG 747  
QY 361 ATGGCACCTGAAGTTTTTGAAGTAGTAATTACAGTGAATAATGTGACGTCTTCAGCTGG 420  
DB 748 ATGGCACCTGAAGTTTTTGAAGTAGTAATTACAGTGAATAATGTGACGTCTTCAGCTGG 807  
QY 421 GGTATTATCTTTGGGAAGTGATAACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 480  
DB 808 GGTATTATCTTTGGGAAGTGATAACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 867  
QY 481 GCTTTCGAATCATGTGGGCTGTTCAATAATGTTACTGACCACTGATAAAAAATTTA 540  
DB 868 GCTTTCGAATCATGTGGGCTGTTCAATAATGTTACTGACCACTGATAAAAAATTTA 927  
QY 541 CCTAAGCCCATTTGAGAGCTGATGACTCGTGTGTTGTTCTTAAGATCCCTCCAGGCGCCT 600  
DB 928 CCTAAGCCCATTTGAGAGCTGATGACTCGTGTGTTGTTCTTAAGATCCCTCCAGGCGCCT 987  
QY 601 TCAATGGAGGAAATGTGAAATAATGACTCACTGATGCGGTACTTTCCAGGAGCAGAT 660  
DB 988 TCAATGGAGGAAATGTGAAATAATGACTCACTGATGCGGTACTTTCCAGGAGCAGAT 1047  
QY 661 GAGCCATTACAGTATCCTTGTCTCAG 684  
DB 1048 GAGCCATTACAGTATCCTTGTCTCAG 1071  
RESULT 9  
AAT85094  
ID AAT85094 standard; cDNA; 2443 BP.  
XX  
AC AAT85094;  
XX  
DT 19-NOV-1997 (first entry)  
XX  
DE Mouse transforming growth factor-beta activated kinase TAK-1 cDNA.  
XX TGF-beta; signal transduction; TGF-beta activated kinase;  
XX MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;  
XX protein kinase; ss.  
XX Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT CDS 157..1896  
FT /\*tag= a  
FT /product= TAK-1  
XX  
XX JP09163990-A.  
XX  
XX 24-JUN-1997.  
XX  
XX 27-SEP-1996; 96JP-0256747.  
XX  
XX 24-JUL-1996; 96US-0685625.  
XX 29-SEP-1995; 95JP-0253549.  
XX  
XX (CHUS ) CHUGAI PHARM CO LTD.  
XX (UENO/) UENO N.  
XX  
XX WPI; 1997-380171/35.  
XX P-PSDB; AAM27092.  
XX  
XX DNA encoding transforming growth factor-beta-activated kinase, TAK-1  
XX - useful for studying the TGF-beta signal transduction system

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 1949.  
XX  
KW Drosophila: developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
(PEKE ) PE CORP NY.  
XX  
Venter JC, Adams M, Li PWD, Myers EW;  
XX  
WPI: 2001-656860/75.  
XX  
P-PSDB; ABB58386.  
XX  
New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Claim 1; SEQ ID NO 1949; 21pp + Sequence Listing; English.  
XX  
The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 3367 BP; 938 A; 898 C; 856 G; 675 T; 0 other;  
Query Match 25.7%; Score 175.6; DB 23; Length 3367;  
Best Local Similarity 56.7%; Pred. No. 8.2e-47;  
Matches 388; Conservative 0; Mismatches 284; Indels 12; Gaps 3;  
Qy 4 GAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTTAATATTGTAAGCTTTATGAGCCTGC 63  
Db 1112 GAGGTGAAGCAGTTGTGCGCGTGAAGCACCCGACATCATCGCTTCGACGGATATCC 1171  
Qy 64 TTGANTC-----CAGTGTGCTGTGTGATGATGATATCTGAAGGGGCTCTTTATATAAT 117  
Db 1172 TCGTACAGCAGGCGCCCTACCTGATATGAGAGTTCGCCGAGGTGATCGCTGCACAC 1231  
Qy 118 GTGCTGATGCTGTGAACCATTCGCCATATTATCTGCTGCCACGCAATGAGTGTGTG 177  
Db 1232 TTCCTTCACGCAAGGTGAAGCGGCATATT---CTCTGCGCCACGCGCATGAGCTGGCG 1288  
Qy 178 TTACAGTGTTCCTCCAGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCAC 237  
Db 1289 CGCCAAATGTGCAGAGGTCTGGCATATTTCATGCCATGACGCCAAACCACTAATCAT 1348  
Qy 238 AGGACCTGAACCAACCAACTTACTGCTGTGTCAGGGGGGACAGTTCTTAAATAATTGT 297  
Db 1349 CGCCAGTGAAGCGCTGACCTGCTCTTGACCAACCAAGGACCGCATCTGAAGATATGC 1408  
Qy 298 GATTTTGTGTACAGCTGTGACATTCAGACACATGACCAATACAGGGAGTGTCTGCT 357  
Db 1409 GACTTCGGCAGCGTGGCGGACAGTGCACCATGATGACCAACAATCGCGGAGTGCCT 1468  
Qy 358 TGGATGGCACTGAGTTTGTGAAGGTAGTATTACAGTGAATAATGTGACGTCTTTCAGC 417

PS Claim 2; Page 10-12; 20pp; Japanese.  
XX  
XX The present sequence encodes mouse transforming growth factor-beta  
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the  
CC TAK-1 protein which is involved in the TGF-beta family signal  
CC transmission system. TAK-1, also known as activator of MAPK Kinase  
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone  
CC morphogenetic protein (BMP) and activates MAPK kinase by  
CC phosphorylation.  
XX  
XX Sequence 2443 BP; 669 A; 567 C; 647 G; 559 T; 1 other;  
Query Match 85.7%; Score 586.4; DB 18; Length 2443;  
Best Local Similarity 91.1%; Pred. No. 3.7e-182;  
Matches 623; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
Qy 1 GTAGAGCTCGGACGATTATCCGCTGTGAACCATCTTAATATTGTAAGCTTTATGAGCC 60  
Db 382 GTGAGCTCGGACGATTGTGCGGTGTGAACCATCTTAATATTGTAAGCTTTATGAGCC 441  
Qy 61 TGCCTGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
Db 442 TGCCTGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 501  
Qy 121 CTGCAATGCTGAACCATGTCATATTATCTGCTGCCACGCAATGATGGTGTGTTA 180  
Db 502 CTGCAATGCTGAACCATGTCATATTATCTGCTGCCACGCAATGATGGTGTGTTA 561  
Qy 181 CAGTGTTCACAGGAGTGTCTATCTTCAGCATGCAACCAAGCGCTTAATTCACAG 240  
Db 562 CAGTGTTCACAGGAGTGTCTATCTTCAGCATGCAACCAAGCGCTTAATTCACAG 621  
Qy 241 GACCTGAACCAACCAACTTACTGCTGTGCTGAGGGGGACAGTCTTAAATAATTGTGAT 300  
Db 622 GACCTGAACCAACCAACTTACTGCTGTGCTGAGGGGGACAGTCTTAAATAATTGTGAT 681  
Qy 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATACAGGGGAGTGTGCTGTG 360  
Db 682 TTTGGTACAGCTGTGACATTCAGACACATGACCAATACAGGGGAGTGTGCTGTG 741  
Qy 361 ATGCGACCTGAAGTTTGTGAAGTAGTAAATTCAGATGAAATAATGTGACGCTTCAGCTGG 420  
Db 742 ATGCGACCTGAAGTTTGTGAAGTAGTAAATTCAGATGAAATAATGTGACGCTTCAGCTGG 801  
Qy 421 GGTATTATCTTTGGGAAGTGATAACCGCTCGAAACCTTTGATGAGATGTTGGTGGCCCA 480  
Db 802 GGTATTATCTTTGGGAAGTGATAACCGCTCGAAACCTTTGATGAGATGTTGGTGGCCCA 861  
Qy 481 GCTTTCGAATCATGTGGCTGTTCATATGTTACTGACACACATGATAAATAATTTA 540  
Db 862 GCTTTCGAATCATGTGGCTGTTCATATGTTACTGACACACATGATAAATAATTTA 921  
Qy 541 CCTAAGCCCATGAGACCTGATGACATGCTGTTGTGTTGTTAAAGATCCTTCCAGCGCCCT 600  
Db 922 CCTAAGCCCATGAGACCTGATGACATGCTGTTGTGTTGTTAAAGATCCTTCCAGCGCCCT 981  
Qy 601 TCAATGAGGAAATTTGAAATAATGACTCATCTGATGCGGTATCTTTCAGAGGAGAT 660  
Db 982 TCAATGAGGAAATTTGAAATAATGACTCATCTGATGCGGTATCTTCCAGAGGAGAT 1041  
Qy 661 GAGCCATTACAGTATCTTGTCTCAG 684  
Db 1042 GAGCCATTACAGTATCTTGTCTCAG 1065  
RESULT 10  
ABL02489 standard; cDNA; 3367 BP.  
XX  
XX ABL02489;  
XX  
XX 26-MAR-2002 (first entry)  
XX





CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 759 BP; 226 A; 162 C; 198 G; 173 T; 0 other;  
Query Match 10.6%; Score 72.2; DB 23; Length 759;  
Best Local Similarity 50.6%; Pred. No. 4.5e-13;  
Matches 201; Conservative 0; Mismatches 193; Indels 3; Gaps 1;  
QY 79 CTGTGATGGAATATCTGTAAGGGGCTCTTTATATAATGTCTGATGCTGCTGAACCA 138  
Db 241 CTGTTGATGAATTCGTAGACGGTGGATCTCTGTCAGTTTCTGCA---CGCGAAAGC 297  
QY 139 TTGCCATATATCTGCTCCACCAAGTGTGTGTTTACAGTGTTCCTCAAGGAGTG 198  
Db 298 AAGCCAAAGTTATTGCGATGCCACCGCTTCAACTGGGCGCATCAGATCGCTCAGGSCATA 357  
QY 199 GCTTATCTTACAGCATGCAACCCAAAGCGCTAATTACAGGGACCTGAAACCCCAAC 258  
Db 358 GCTATCTCCATGCGATGCGGAAAGCAGTAATTCATCGCATATAAGCCACTCAAT 417  
QY 259 TTACTGCTGTTGCGAGGGGCGACGTTCTPAAAAATTGTGATTTGGTACAGCTGTGAC 318  
Db 418 AACTGCTATGCGAGAGGACTCAAACTGAAGATTGGGATTTGCGAACTGTTGTGGAC 477  
QY 319 ATTACACACATGACCAATACAAAGGGAGTGTCTTGGATGGACCTGAAGTTT 378  
Db 478 CTATCCCAATCGATATCGTCAATGCGGCGACCTCGAGATACAAAGCGCGGAGTAAG 537  
QY 379 GAAGGTAGTAATACAGTGAATAATGTGAGCTTTCAGCTGGGGTATTATTCTTTGGGAA 438  
Db 538 GAGTTGTTGATTTCAATCCATCGAATAATCAACCAACCGCGCTTTCAAAG 597  
QY 439 GTGATAACGCGTCGGAACCCCTTTGATGAGATTGGTG 475  
Db 598 GTTCTACAGGAATAAACCCGATGAAAGTGGGATG 634  
RESULT 13  
ABZ13077  
ID ABZ13077 standard; DNA; 1428 BP.  
AC ABZ13077;  
XX  
XX 21-JAN-2003 (first entry)  
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 882.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX Arabidopsis thaliana.  
XX WO200216655-A2.  
XX 28-FEB-2002.  
XX 24-AUG-2001; 2001WO-US26685.  
XX 24-AUG-2000; 2000US-227866P.  
XX 26-JAN-2001; 2001US-264647P.  
XX 22-JUN-2001; 2001US-300111P.  
XX (SRI ) SCRIPPS RES INST.  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.  
XX Identifying a stress condition to which a plant cell has been exposed  
XX and producing plants with increased tolerance to these abiotic stresses  
XX -  
XX Claim 144; SEQ ID NO 882; 577pp + Sequence Listing; English.  
XX PS

XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising:  
CC (a) contacting nucleic acid representative of expressed polynucleotides  
CC in the plant cell with an array or probes representative of the plant  
CC cell genome; and  
CC (b) detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.  
XX  
SQ Sequence 1428 BP; 432 A; 257 C; 342 G; 397 T; 0 other;  
Query Match 9.6%; Score 65.4; DB 24; Length 1428;  
Best Local Similarity 49.3%; Pred. No. 1.1e-10;  
Matches 171; Conservative 0; Mismatches 176; Indels 0; Gaps 0;  
QY 338 ATAAACAGGGAGTGTCTGCTGATGGCAGCTGAGTTTGTAGGTAGTAAATTACAGTG 397  
Db 956 ATAACATAGGAACCTTATAGGTGATGGCAGCTGAGTTTAAACCGATACCATGGAC 1015  
QY 398 AAAAATGTGACGCTTTCAGCTGGGGTATTATTCTTTGGGAAGTGAATACGCGTCGGAAC 457  
Db 1016 GGAAGTGGCATGTTTATAGTTTGGACTTCTTTATGGGAATGGTAGTGAGCACTTC 1075  
QY 458 CCTTTGATGAGATGCTGCGCCAGCTTCCGAATCATGTGGGCTGTTCATATGGTACTC 517  
Db 1076 CATATGAGGAGATGAATTTGCTGAACAAATTCCTACGAGTTATATACAGAAAATTA 1135  
QY 518 GACCACCACTGATAAAAAATTACCTAAGCCATTGAGCCCTGATGACCTCGTTGTTGTT 577  
Db 1136 GGCCAGTTATACCGACGGATTGTCCAGCGGCCCAATGAAGAGCTGTGACCGGATGTTG 1195  
QY 578 CTAAAGATCTTCCAGCGCCCTTCAATGAGGAAAATTGTGAAAATAATGACTCACTTGA 637  
Db 1196 CATCGCAACAGACAGACGACCGGAATCTGCGAGATTGTCAAAGTGTGGAACATTTC 1255  
QY 638 TGCAGTACTTCCAGGAGCAGATGAGCCATTACAGTATCCTTGTCTAG 684  
Db 1256 AGAAGTCTCTACAAGCGAAGGAAACTTAATCTTTTACCTAGCCAG 1302  
RESULT 14  
ABZ12828  
ID ABZ12828 standard; DNA; 1398 BP.  
AC ABZ12828;  
XX  
XX 21-JAN-2003 (first entry)  
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 633.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX Arabidopsis thaliana.  
XX WO200216655-A2.  
XX 28-FEB-2002.  
XX 24-AUG-2001; 2001WO-US26685.  
XX 24-AUG-2000; 2000US-227866P.  
XX 26-JAN-2001; 2001US-264647P.  
XX 22-JUN-2001; 2001US-300111P.  
XX (SRI ) SCRIPPS RES INST.  
XX PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.  
XX Identifying a stress condition to which a plant cell has been exposed  
XX and producing plants with increased tolerance to these abiotic stresses  
XX -  
XX Claim 144; SEQ ID NO 882; 577pp + Sequence Listing; English.  
XX PS

```

XX Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed
XX and producing plants with increased tolerance to these abiotic stresses
XX
XX Claim 144; SEQ ID NO 633; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising:
XX (a) contacting nucleic acid representative of expressed polynucleotides
XX in the plant cell with an array or probes representative of the plant
XX cell genome; and
XX (b) detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention.
XX
XX Note: The sequence data for this patent is not represented in the printed
XX specification but is based on sequence information supplied to Derwent by
XX the European Patent Office.
XX
XX Sequence 1398 BP; 420 A; 284 C; 338 G; 356 T; 0 other;
XX
XX
XX Query March 8.3%; Score 56.6; DB 24; Length 1398;
XX Best Local Similarity 52.3%; Pred. NO. 8.3e-08;
XX Matches 150; Conservative 0; Mismatches 134; Indels 3; Gaps 1;
XX
XX QY 186 TTCCCAAGGAGTGGCTTATCTTCACGATGCAACCCCAAGCGCTAATTCACAGGACCT 245
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX
XX QY 246 GAAACCAACCAACTTACTGCTGGTTGAGGGGGGACAG---TTCTAAAAATTTGTTATTT 302
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX
XX Db 774 GAAACCGGAAACATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 833
XX
XX QY 303 TGGTACAGCTGTGACATTGACACACATGACCAATTAACAGGGGAGTGTGCTTGGAT 362
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX
XX QY 363 GGCACTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX
XX QY 423 TATTATTTTGGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 469
XX Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX
XX Db 954 GATTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1000
XX
XX
XX RESULT 15
XX ID AAC45777
XX AC AAC45777;
XX
XX DT 18-OCT-2000 (first entry)
XX
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47724.
XX
XX KW Hybridization assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX
XX OS Arabidopsis thaliana.
XX
XX PN EP1033405-A2.
XX
XX XX 06-SEP-2000.
XX
XX

```

```

PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.

```

PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.

PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 8.3%; Score 56.6; DB 21; Length 1401;  
Best Local Similarity 52.3%; Pred. No. 8.3e-08;  
Matches 150; Conservative 0; Mismatches 134; Indels 3; Gaps 1;  
Qy 186 TTCCCAAGGAGTGGCTTATCTTCACAGCATGGAACCCAAAGCGCTTATTCACAGGACCT 245  
Db 714 TGCAGAAATCGTCTCTGCAGTTTTCCCATCTCCATGAGAAAGGATATGACAGAGATCT 773  
Qy 246 GARACCAACCAAACTTACTGCTGTGTCAGGGGGACAG---TTCTAAAAATTTTGATTT 302  
Db 774 GAAACCCGAAACATCTCTATGATACAGATGCCACGTGATGCTAACAGATTTTGTTT 833  
Qy 303 TGGTACAGCCTGTGCATTCAGACATTCAGACACATGACCAATAACAGGGGAGTCTGCTTGGAT 362  
Db 834 AGCAAGGAATTTGAGAAACACACAGATCAAATCCATGTGCGGAACCTACGGAGTATAT 893  
Qy 363 GGCACCTGAAGTTTTTGAAGGTAGTATTACAGTGAATAATGTACCTTTCAGCTGGG 422  
Db 894 GGCACCTGAATTTGTTTCAGAGAAAGACATGATAAAGCAGCTGACTGGTGGAGCGTAGG 953  
Qy 423 TATTATCTTTGGGAAGTGATACGGCTCGGAAACCCCTTTTGATGAGA 469  
Db 954 GATTCCTCTGTATGATGATCTCAGGAAAGCCACCGCTTTCGGGA 1000

Search completed: December 4, 2003, 05:02:26  
Job time : 357.013 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 00:51:52 ; Search time 3168.12 Seconds  
(without alignments)  
8832.421 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091

Perfect score: 684

Sequence: 1 gttagagcttcggcagttatc.....cattacagatcctgtgcag 684

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pi:\*

26: em\_ro:\*

27: em\_sts:\*

28: em\_un:\*

29: em\_vi:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rnd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	684	100.0	1704	6	E38399	E38399 NF-kappa B
2	684	100.0	1705	9	AB009358	AB009358 Homo sapi
3	684	100.0	1745	9	AF218074	AF218074 Homo sapi
4	684	100.0	1788	6	AR231198	AR231198 Sequence
5	684	100.0	1788	6	AR307979	AR307979 Sequence
6	684	100.0	2656	6	AR231192	AR231192 Sequence
7	684	100.0	2656	6	AR307973	AR307973 Sequence
8	684	100.0	2757	9	BC017715	BC017715 Homo sapi
9	684	100.0	2769	6	AX377912	AX377912 Sequence
10	684	100.0	2769	6	AB009356	AB009356 Homo sapi
11	684	100.0	2785	6	E38397	E38397 NF-kappa B
12	684	100.0	2850	9	AB009357	AB009357 Homo sapi
13	684	100.0	2866	6	E38398	E38398 NF-kappa B
14	586.4	85.7	2443	10	MUSTAK1	D76446 Mouse mRNA
15	586.4	85.7	3107	10	BC006665	BC006665 Mus muscu
16	523.6	76.5	135308	2	AC114407	AC114407 Mus muscu
17	498.4	72.9	2745	5	BC049005	BC049005 Xenopus lae
18	498.4	72.9	2812	5	XL092030	U92030 Xenopus lae
19	177.2	25.9	3349	3	AF199466	AF199466 Drosophil
20	175.6	25.7	3386	3	AY051953	AY051953 Drosophil
21	144.8	21.2	153168	2	AC025387	AC025387 Homo sapi
22	141.2	20.6	100994	9	HSJ154G14	AL121964 Human DNA
23	141.2	20.6	200469	2	AL161435	AL161435 Homo sapi
24	140.2	20.5	153168	2	AC025387	AC025387 Homo sapi
25	140.2	20.5	200469	2	AL161435	AL161435 Homo sapi
26	138	20.2	3482	9	HSM800550	AL050393 Homo sapi
27	127	18.6	144836	2	AC011102	AC011102 Homo sapi
28	124.6	18.2	165700	10	AL833781	AL833781 Mouse DNA
29	124.6	18.2	233722	2	AC134067	AC134067 Rattus no
30	124.6	18.2	238817	2	AC134221	AC134221 Rattus no
31	124.6	18.2	242458	2	AC094220	AC094220 Rattus no
32	108	15.8	32564	2	AC014558	AC014558 Drosophil
33	108	15.8	164942	3	AC011758	AC011758 Drosophil
34	108	15.8	302540	3	AE003571	AE003571 Drosophil
35	97.4	14.2	174275	5	EX119992	EX119992 Zebrafish
36	97.4	14.2	192888	2	EX005296	EX005296 Danio rer
37	65.8	9.6	1611	3	AF145377	AF145377 Caenorhab
38	65.8	9.6	1993	3	AF143242	AF143242 Caenorhab
39	65.4	9.6	1428	6	AX506187	AX506187 Sequence
40	65.4	9.6	1428	6	AX651822	AX651822 Sequence
41	65.2	9.5	14603	3	AC116959	AC116959 Dictyoste
42	64.6	9.4	150133	2	AC014084	AC014084 Drosophil
43	64.6	9.4	186607	3	AC007772	AC007772 Drosophil
44	64.6	9.4	197071	3	AC007771	AC007771 Drosophil
45	64.6	9.4	224529	3	AE003732	AE003732 Drosophil

ALIGNMENTS

RESULT 1  
E38399  
LOCUS E38399 1704 bp DNA linear PAT 31-JAN-2002  
DEFINITION NF-kappa B activation inhibitory drug targeting TAK1 and method for  
identifying the same.  
ACCESSION E38399  
VERSION E38399.1 GI:18626979  
KEYWORDS JP 2000197500-A/5.  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1704)  
AUTHORS Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.  
TITLE NF-kappa B activation inhibitory drug targeting TAK1 and method for  
identifying the same





[illegible]

QY	Db	661	GAGCCATTACAGTATCCTTGTGAG	684
		892	GAGCCATTACAGTATCCTTGTGAG	915
RESULT 4				
AR231198				
LOCUS		AR231198	1788 bp	DNA
DEFINITION		Sequence 14 from patent US 6451617.		linear
ACCESSION		AR231198		
VERSION		AR231198.1	GI:27272086	
KEYWORDS				
SOURCE		Unknown.		
ORGANISM		Unknown.		
REFERENCE		1 (bases 1 to 1788)		
AUTHORS		Ono, K., Ohtomo, T. and Tsuchiya, M.		
TITLE		Method of screening TGF-beta inhibitory substances		
JOURNAL		Patent: US 6451617-A 14 17-SEP-2002;		
FEATURES		Location/Qualifiers		
		1..1788		
source		/organism="unknown"		
BASE COUNT		551 a	394 c	413 g
ORIGIN		430 t		
		Query Match	100.0%;	Score 684; DB 6; Length 1788;
		Best Local Similarity	100.0%;	Pred. No. 2e-200;
		Matches 684; Conservative	0; Mismatches	0; Indels 0; Gaps 0;
QY		1	GTAGAGCTTCGGCAGTGTATCCCGTGTGAACCATCCTAATATGTAAAGCTTTATGGAGCC	60
Db		232	GTAGAGCTTCGGCAGTGTATCCCGTGTGAACCATCCTAATATGTAAAGCTTTATGGAGCC	291
QY		61	TGCTTGAATCCAGTGTGCTTTGTGTATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG	120
Db		292	TGCTTGAATCCAGTGTGCTTTGTGTATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG	351
QY		121	CTGCATGTGTGTGAACCAATGCCATATTATATCTGTCGCCACGCAATGAGTTGGTGTGTTTA	180
Db		352	CTGCATGTGTGTGAACCAATGCCATATTATATCTGTCGCCACGCAATGAGTTGGTGTGTTTA	411
QY		181	CAGTGTTCCTCCAGAGTGGCTTATCTTCACAGCATGCAACCCCAAGCGCTAATTACACAGG	240
Db		412	CAGTGTTCCTCCAGAGTGGCTTATCTTCACAGCATGCAACCCCAAGCGCTAATTACACAGG	471
QY		241	GACCTGAAACCAACCAAACTTACTGCTGGTGTGCAGGGGGGACAGTTCTAAAAATTTGTGAT	300
Db		472	GACCTGAAACCAACCAAACTTACTGCTGGTGTGCAGGGGGGACAGTTCTAAAAATTTGTGAT	531
QY		301	TTTGGTACAGCTGTGCATTTACAGACACACATGACACATTAACAGGGGAGTGTCTCTTGG	360
Db		532	TTTGGTACAGCTGTGCATTTACAGACACACATGACACATTAACAGGGGAGTGTCTCTTGG	591
QY		361	ATGGCACTGTAGTATTTTGAAGGTAGTAATACAGTGAATAATGTGAGCTCTTCAGCTGG	420
Db		592	ATGGCACTGTAGTATTTTGAAGGTAGTAATACAGTGAATAATGTGAGCTCTTCAGCTGG	651
QY		421	GGTATATTCTTTTGGGAAGTATAACGGCTGTGGAAACCCCTTTGATGAGATTGGTGGCCCA	480
Db		652	GGTATATTCTTTTGGGAAGTATAACGGCTGTGGAAACCCCTTTGATGAGATTGGTGGCCCA	711
QY		481	GCTTTCCGAATCATGTGGCTGTTTCAATATGTACTCGACACCACTGATAAAAAATTTTA	540
Db		712	GCTTTCCGAATCATGTGGCTGTTTCAATATGTACTCGACACCACTGATAAAAAATTTTA	771
QY		541	CCTAAGCCCATTTGAGAGCCTGTATGACTCCGTTGTGGTCTTAAAGATCCCTTCCAGCGCCCT	600
Db		772	CCTAAGCCCATTTGAGAGCCTGTATGACTCCGTTGTGGTCTTAAAGATCCCTTCCAGCGCCCT	831
QY		601	TCAATGGAGAAATTTGTGAAAATAATGACTCACTTGATCGGTACTTTCCAGGACGAGAT	660
Db		832	TCAATGGAGAAATTTGTGAAAATAATGACTCACTTGATCGGTACTTTCCAGGACGAGAT	891



[illegible]

```
QY 661 GAGCATTACAGTATCCCTGTGAC 684
DB 1068 GAGCATTACAGTATCCCTGTGAC 1091

RESULT 8
BC017715 2757 bp mRNA linear PRI 06-DEC-2001
LOCUS Homo sapiens, mitogen-activated protein kinase kinase 7,
DEFINITION clone MGC:21263 IMAGE:3906837, mRNA, complete cds.
ACCESSION BC017715
VERSION BC017715.1 GI:17389342
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL Strausberg, R.
TITLE Direct Submission
INSTITUTE National Institutes of Health, Mammalian
INSTITUTE Gene Collection (MGC), Cancer Genomics Office, National Cancer
INSTITUTE Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
INSTITUTE USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Contact: (Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 22 Row: 1 Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4507360.
Location/Qualifiers
1. .2757
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:6885"
/db_xref="taxon:9606"
/clone="MGC:21263 IMAGE:3906837"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
198..1937
/codon_start=1
/product="mitogen-activated protein kinase kinase
7"
/protein_id="AAH17715.1"
/db_xref="GI:17389343"
/translation="MSTVSAASSSSSSAGEMIEAPSOVLNFBFIDYKEIEVEVVR
GAGVVCCKAKRAKVAIKOIESESRKAFIVLRQLSRVNHPIVKLYGACLNVPCL
VMEYAEGLSLYNVLHGAEPIPVYTAHAMSKLQCSOGVAYLHSMOPKALIHRLKPP
NLLVAGTGVKICDFGAGTACIQTHMNNKSGSAWAPVPEGSSNYSEKCDVSGI
LWVITRKPDEITGGFAFRIMVAHNGTAPPLIKLPKEISLMKNCWKSQSPRS
MEVIMKTHLMRYPPGADEPFOYQYDEGQSNATSGSFMDSASTNKSQSDTN
MEQVPAINDITKIKESLKNQAKQSGESGLSLGSGSVESLPTSGKMSADM
SETPARTADTDNGQPRRSIQDLTVTGTPGVSSRSRSPSVMTTSGTSEKPR
SHKTPDDSDTDNGSIPMAYITLDHQLQPLAPCPNSKESMAVFPQHKMAQVYK
VQIEALLQKQKQELVQLDQKQNTSELVQEHKLLDENKSLSTYYQQCKKQLE
VIRSQQKQKQTS"
BASE COUNT 819 a 566 c 647 g 725 t
ORIGIN
```

```
Query Match 100.0%; Score 684; DB 9; Length 2757;
Best Local Similarity 100.0%; Pred. No. 1.9e-200;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAGAGCTTCGGCAGTTATCCGTTGTAACCATCTTAATATTGTAAAGCTTTATGAGCC 60
DB 423 GTAGAGCTTCGGCAGTTATCCGTTGTAACCATCTTAATATTGTAAAGCTTTATGAGCC 482
QY 61 TGCATTGAATCAGTGTCTTGTGTGATGAATATGCTGAAGGGGGCTCTTTATATAATGTS 120
DB 483 TGCATTGAATCAGTGTCTTGTGTGATGAATATGCTGAAGGGGGCTCTTTATATAATGTS 542
QY 121 CTGCAATGGTGTGAACCAATTCATATTAATCTGCTGCCCCACGCAATGAGTTGGTGTTA 180
DB 543 CTGCAATGGTGTGAACCAATTCATATTAATCTGCTGCCCCACGCAATGAGTTGGTGTTA 602
QY 181 CAGTGTTCCTCAAGAGAGTGTCTTATCTTACAGCATGCAACCAAGCGCTAATTCACAGG 240
DB 603 CAGTGTTCCTCAAGAGAGTGTCTTATCTTACAGCATGCAACCAAGCGCTAATTCACAGG 562
QY 241 GACCTGGAACCAACCAAACTTACTGCTGTTGAGGGGGGACAGTTCTTAAAAATTTGTGAT 300
DB 663 GACCTGGAACCAACCAAACTTACTGCTGTTGAGGGGGGACAGTTCTTAAAAATTTGTGAT 722
QY 301 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTGCTGG 360
DB 723 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTGCTGG 782
QY 361 ATGGCAGCTGAAGTTTTCAGAGTAGTAAATTAACAGTGAAGAAATGACCTCTTCAGCTGG 420
DB 783 ATGGCAGCTGAAGTTTTCAGAGTAGTAAATTAACAGTGAAGAAATGACCTCTTCAGCTGG 842
QY 421 GGTATTATTCTTTGGGAAGTGATAACGCTCGGAAACCCCTTTGATGAGATTGTTGGCCCA 480
DB 843 GGTATTATTCTTTGGGAAGTGATAACGCTCGGAAACCCCTTTGATGAGATTGTTGGCCCA 902
QY 481 GCTTCCGGAATCATGTGGGCTGTTCAATATGTTGACCTGACCTGATGCGGAGTGTTCAGGAGGAGAT 540
DB 903 GCTTCCGGAATCATGTGGGCTGTTCAATATGTTGACCTGACCTGATGCGGAGTGTTCAGGAGGAGAT 962
QY 541 CCTAACCCATTGAGAGCCTGATGACCTGCTGTTGCTGTAAAGATCCTTCCAGAGCCCT 600
DB 963 CCTAACCCATTGAGAGCCTGATGACCTGCTGTTGCTGTAAAGATCCTTCCAGAGCCCT 1022
QY 601 TCAATGGAGAAATCTGAAATAATGACCTCACTGATGCGGTACTTCAGGAGGAGGAGAT 660
DB 1023 TCAATGGAGAAATCTGAAATAATGACCTCACTGATGCGGTACTTCAGGAGGAGGAGAT 1082
QY 661 GAGCCATTACAGTATCCTTGTGAC 684
DB 1083 GAGCCATTACAGTATCCTTGTGAC 1106

RESULT 9
AX377912 2769 bp DNA linear PAT 18-MAR-2002
LOCUS Sequence 107 from Patent WO0212338.
DEFINITION AX377912
ACCESSION AX377912
VERSION AX377912.1 GI:19573976
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL Gillen, C., Wetzel, I., Wnendt, S., Weihe, E. and Schaefer, M.K.
TITLE Screening method
PATENT Patent: WO 0212338-A 107 14-FEB-2002;
JOURNAL Gruenenthal GmbH (DE)
FEATURES Location/Qualifiers
1..2769
/organism="Homo sapiens"
/mol_type="genomic DNA"
```

```
BASE COUNT      811 a 565 c 640 g 753 t
ORIGIN
Query Match      100.0%; Score 684; DB 6; Length 2769;
Best Local Similarity 100.0%; Pred. No. 1.9e-200;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTATATATTGTAAGCTTTATGAGCC 60
Db 388 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTATATATTGTAAGCTTTATGAGCC 447
QY 51 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATG 120
Db 448 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATG 507
QY 121 CTGATGTGTGACCAATTCGCAATATATATATGCTGCGCCAGCAATGAGTGTGTTA 180
Db 508 CTGATGTGTGACCAATTCGCAATATATATGCTGCGCCAGCAATGAGTGTGTTA 567
QY 181 CAGTGTTCCTCAAGGAGTGGCTTATCTTCAGCATGCAACCCAAAGCGCTAATTCACAG 240
Db 568 CAGTGTTCCTCAAGGAGTGGCTTATCTTCAGCATGCAACCCAAAGCGCTAATTCACAG 627
QY 241 GACCTGAACCCACCAAACTTACTGCTGAGGGGGAGCTTCTTAAATTTTGTAT 300
Db 628 GACCTGAACCCACCAAACTTACTGCTGAGGGGGAGCTTCTTAAATTTTGTAT 687
QY 301 TTTGGTACAGCTGTGACATTCAGACATGACCAATTAACCAAGGGGAGTGTGCTTGG 360
Db 688 TTTGGTACAGCTGTGACATTCAGACATGACCAATTAACCAAGGGGAGTGTGCTTGG 747
QY 361 ATGGCACTGAAAGTTTGTGAAGGTAGTAATACAGTGAATAATGAGCTTCTGAGTGG 420
Db 748 ATGGCACTGAAAGTTTGTGAAGGTAGTAATACAGTGAATAATGAGCTTCTGAGTGG 807
QY 421 GGTATTATCTTTGGGAAGTGAATACCGCTCGGAAACCCCTTTGATGAGATTGGGCCA 480
Db 808 GGTATTATCTTTGGGAAGTGAATACCGCTCGGAAACCCCTTTGATGAGATTGGGCCA 867
QY 481 GCTTTCGAATCATGTGGCTGTTTCAATATGTTACTGACCACTGATAAAAAATTTA 540
Db 868 GCTTTCGAATCATGTGGCTGTTTCAATATGTTACTGACCACTGATAAAAAATTTA 927
QY 541 CCTAAGCCCATGTGAGCCTGATCACTCGTTGTTGTTTAAAGATCTTCCAGCGCCT 600
Db 928 CCTAAGCCCATGTGAGCCTGATCACTCGTTGTTGTTTAAAGATCTTCCAGCGCCT 987
QY 601 TCAATGAGAAATTTGAAATAATGACCTCACTTGTGCGGTACTTTCCAGGAGCAGAT 660
Db 988 TCAATGAGAAATTTGAAATAATGACCTCACTTGTGCGGTACTTTCCAGGAGCAGAT 1047
QY 661 GAGCATTACAGTATCCTCTCAG 684
Db 1048 GAGCATTACAGTATCCTCTCAG 1071

RESULT 10
AB009356 2769 bp mRNA linear PRI 04-MAR-1998
LOCUS Homo sapiens mRNA for TGF-beta activated kinase 1a, complete cds.
DEFINITION Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ACCESSION AB009356.1 GI:2924623
VERSION TAK1a; TGF-beta activated kinase 1a.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE Sakurai, H., Shigemori, N., Hasegawa, K. and Sugita, T.
AUTHORS TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
TITLE NF-kappa B-inducing kinase-independent mechanism
JOURNAL Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
```

MEDLINE 98153801  
PUBMED 9480845  
REFERENCE 2 (bases 1 to 2769)  
AUTHORS Sakurai, H.  
TITLE Direct Submission  
JOURNAL Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research laboratory, Tanabe Seliyaku Co. Ltd.; 16-89, Kashima-3-chome, Yodogawa-Ku, Osaka, Osaka 532, Japan (E-mail:hsakurai@tanabe.co.jp, Tel:++81 6 300 2571, Fax:++81 6 300 2593)

FEATURES  
Location/Qualifiers  
1..2769  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="pBSTAK1a"  
/tissue\_type="lung"  
/clone\_lib="Lambda gt11"  
163..1902  
/note="TAK1a"  
/codon\_start=1  
/product="TGF-beta activated kinase 1a"  
/protein\_id="BAA25025.1"  
/db\_xref="GI:2924624"

CDS  
/translation="MSTASASSSSSSSAGEMIEAPSVQLNPFEDIDYKEIEVEEVVGR  
GAFGVCAKRWKADVAIKQIESERKAFIVELQSRVNHPIVKLYGLACLNVPCL  
VMEYAGGSLYNVLHGAEPLEYTAHAHSMCLQCSQGVAYLHMQPALKALHDLKPP  
LLJLVAGGTVLKI CDFTGACDIOTMTNNGSAAMAPEVPEGSNYSKEDVFSWGI  
LMETVTRKPFDEIGGPAFRIMMAVHNGTRPLIKLPKPIESLMTRCWSKDSORPS  
MEIVKIMTHLMRYFPGADEPLOYCOYSDGQSNATSGSPMDIASNTNSKSDTN  
MEQVPATNDITKRLSKLLKNOAQQESERLSLGSAGRSVSVEPPTSEGRWSADM  
SEIEARLAATGTGQPRRSIQDLTVGTGEPQVSSRSSPSVEMITTSSEKRPTR  
SHMTPDDSDTSDNSIMPAVLTLDHLOPLAPCPNSKESMAVFPQHKMAQBYMK  
VQTEIALLLQKQLVAELQDEKQNTSRLVQEHKLLDENKLSLSTYYQQCKQLF  
VIRSQQKRGQTS"

BASE COUNT 811 a 565 c 640 g 753 t  
ORIGIN

Query Match 100.0%; Score 684; DB 9; Length 2769;  
Best Local Similarity 100.0%; Pred. No. 1.9e-200;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTATATTTGTAAGCTTTATGAGCC 60  
Db 388 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTATATTTGTAAGCTTTATGAGCC 447  
QY 61 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATG 120  
Db 448 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATG 507  
QY 121 CTGATGTGTGCAACCATTCGCAATATATATATGCTGCTCCACGCAATGAGTGTGTTA 180  
Db 508 CTGATGTGTGCAACCATTCGCAATATATATGCTGCTCCACGCAATGAGTGTGTTA 567  
QY 181 CAGTGTTCCTCAAGGAGTGGCTTATCTTCAGCATGCAACCCAAAGCGCTAATTCACAG 240  
Db 568 CAGTGTTCCTCAAGGAGTGGCTTATCTTCAGCATGCAACCCAAAGCGCTAATTCACAG 627  
QY 241 GACCTGAACCCACCAAACTTACTGCTGAGGGGGAGCTTCTTAAATTTTGTAT 300  
Db 628 GACCTGAACCCACCAAACTTACTGCTGAGGGGGAGCTTCTTAAATTTTGTAT 687  
QY 301 TTTGGTACAGCTGTGACATTCAGACATGACCAATTAACCAAGGGGAGTGTGCTTGG 360  
Db 688 TTTGGTACAGCTGTGACATTCAGACATGACCAATTAACCAAGGGGAGTGTGCTTGG 747  
QY 361 ATGGCACTGAAAGTTTGTGAAGGTAGTAATACAGTGAATAATGAGCTTCTGAGTGG 420  
Db 748 ATGGCACTGAAAGTTTGTGAAGGTAGTAATACAGTGAATAATGAGCTTCTGAGTGG 807  
QY 421 GGTATTATCTTTGGGAAGTGAATACCGCTCGGAAACCCCTTTGATGAGATTGGGCCA 480  
Db 808 GGTATTATCTTTGGGAAGTGAATACCGCTCGGAAACCCCTTTGATGAGATTGGGCCA 867  
QY 481 GCTTTCGAATCATGTGGCTGTTTCAATATGTTACTGACCACTGATAAAAAATTTA 540  
Db 868 GCTTTCGAATCATGTGGCTGTTTCAATATGTTACTGACCACTGATAAAAAATTTA 927  
QY 541 CCTAAGCCCATGTGAGCCTGATCACTCGTTGTTGTTTAAAGATCTTCCAGCGCCT 600  
Db 928 CCTAAGCCCATGTGAGCCTGATCACTCGTTGTTGTTTAAAGATCTTCCAGCGCCT 987  
QY 601 TCAATGAGAAATTTGAAATAATGACCTCACTTGTGCGGTACTTTCCAGGAGCAGAT 660  
Db 988 TCAATGAGAAATTTGAAATAATGACCTCACTTGTGCGGTACTTTCCAGGAGCAGAT 1047  
QY 661 GAGCATTACAGTATCCTCTCAG 684  
Db 1048 GAGCATTACAGTATCCTCTCAG 1071

```

481 GCTTTCGGAATCATGTGGGCTGTTTCATTAATGTGACTCGACCACTGATATAAAATTTA 540
    |||
868 GCTTTCGGAATCATGTGGGCTGTTTCATTAATGTGACTCGACCACTGATATAAAATTTA 927
    |||
541 CTTAGCCCATGAGAGCTGATGCTGCTGTTGCTTAAAGATCTTCCAGGCGCCCT 600
    |||
928 CTTAGCCCATGAGAGCTGATGCTGCTGTTGCTTAAAGATCTTCCAGGCGCCCT 987
    |||
601 TCAATGGAGGAATTTGGAATAAATGACTCACTTGAATCGGTACTTCCAGAGCAGAT 660
    |||
988 TCAATGGAGGAATTTGGAATAAATGACTCACTTGAATCGGTACTTCCAGAGCAGAT 1047
    |||
661 GAGCCATTACAGTATCCTTGTCTAG 684
    |||
1048 GAGCCATTACAGTATCCTTGTCTAG 1071
    |||

RESULT 11
E38397
LOCUS
DEFINITION
NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same.
ACCESSION
E38397
VERSION
E38397.1 GI:18626977
KEYWORDS
JP 2000197500-A/3.
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 2785)
AUTHORS
Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.
TITLE
NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same
JOURNAL
Patent: JP 2000197500-A 3 18-JUL-2000;
TANABE SEIYAKU CO LTD
COMMENT
OS Unidentified
PN JP 2000197500-A/3
PD 18-JUL-2000
PF 04-FEB-1999 JP 1999026803
PI NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI
HASEGAWA
PC C12Q1/48,A61K31/00,A61K45/00,C12N5/10,C12N9/99,PC
C12Q1/02,
PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09,
C12R1:91),
PC C12N5/00,C12N15/00,(C12N15/00,C12R1:91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..2785
FT Location/Qualifiers
1..2785
/organism="Unidentified".
FEATURES
source
1..2785
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 827 a 565 c 640 g 753 t
ORIGIN
Query Match 100.0%; Score 684; DB 6; Length 2785;
Best Local Similarity 100.0%; Pred. No. 1.9e+200;
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAGAGCTCGGCAGTATCCGCTGTGAACCATCTTAATGTAAAGCTTTATGAGCC 60
DB 388 GTAGAGCTTCGGCAGTATCCGCTGTGAACCATCTTAATGTAAAGCTTTATGAGCC 447
QY 61 TCCTTGAATCAGTGTGCTGCTGATGATATGCTGAGGGGCTCTTTATATAATGTG 120
DB 448 TCGTTGAATCAGTGTGCTGCTGATGATATGCTGAGGGGCTCTTTATATAATGTG 507
QY 121 CTGATGCTGTGAACCATTTGCCATTTATATCTGCTGCCAGCAATGAGTTGGTTTA 180
    |||

```

```

Db 508 CTGCATGGTGTGAACCACTTGCCATTTATATCTGCTGCCACGCAATGAGTTGGTTTA 567
QY 181 CAGTGTTCCTCCAGAGTGTCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAG 240
    |||
568 CAGTGTTCCTCCAGAGTGTCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAG 627
    |||
241 GACCTGAAACACCAACCAACTTACTGCTGTTGCAGGGGGACAGTCTTAAAAATTTGTGAT 300
    |||
628 GACCTGAAACACCAACCAACTTACTGCTGTTGCAGGGGGACAGTCTTAAAAATTTGTGAT 687
    |||
301 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATAACAAGGGAGTGTCTGTTG 360
    |||
688 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATAACAAGGGAGTGTCTGTTG 747
    |||
361 ATGGCACCTGAAGTTTTCGAGGTAGTAAATACAGTGAAAAATGACCTCTTCAGCTG 420
    |||
748 ATGGCACCTGAAGTTTTCGAGGTAGTAAATACAGTGAAAAATGACCTCTTCAGCTG 807
    |||
421 GGTATTATTCTTTGGGAAGTATAACGCTGCGAAACCCCTTTGATGAGATTGGTGGCCCA 480
    |||
808 GGTATTATTCTTTGGGAAGTATAACGCTGCGAAACCCCTTTGATGAGATTGGTGGCCCA 867
    |||
481 GCTTTCCGAATCATGTGGCTGTTCATTAATGCTACTCGACCCACCTGATAAAAAATTTA 540
    |||
868 GCTTTCCGAATCATGTGGCTGTTCATTAATGCTACTCGACCCACCTGATAAAAAATTTA 927
    |||
541 CCTAAGCCCATGAGAGCTGATGACTCGCTGTTGCTTAAAGATCCTTCCACAGCGCCCT 600
    |||
928 CCTAAGCCCATGAGAGCTGATGACTCGCTGTTGCTTAAAGATCCTTCCACAGCGCCCT 987
    |||
601 TCAATGGAGGAATTTGGAATAAATGACTCACTTGAATCGGTACTTCCAGAGCAGAT 660
    |||
988 TCAATGGAGGAATTTGGAATAAATGACTCACTTGAATCGGTACTTCCAGAGCAGAT 1047
    |||
661 GAGCCATTACAGTATCCTTGTCTAG 684
    |||
1048 GAGCCATTACAGTATCCTTGTCTAG 1071
    |||

RESULT 12
AB009357
LOCUS
DEFINITION
Homo sapiens mRNA for TGF-beta activated kinase 1b, complete cds.
ACCESSION
AB009357
VERSION
AB009357.1 GI:2924625
KEYWORDS
TAK1b; TGF-beta activated kinase 1b.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (sites)
AUTHORS
Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE
TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
NF-kappa B-inducing kinase-independent mechanism
JOURNAL
Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
MEDLINE
98153801
PUBMED
9480845
REFERENCE
2 (bases 1 to 2850)
AUTHORS
Sakurai,H.
DIRECT SUBMISSION
Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research
laboratory, Tanabe Seiyaku Co. Ltd.; 16-89, Kashima-3-chome,
Yodogawa-ku, Osaka, Osaka 532, Japan (E-mail:hsakurai@tanabe.co.jp.
Tel:++81 6 300 2571, Fax:++81 6 300 2593)
FEATURES
source
Location/Qualifiers
1..2850
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pBSTAK1b"
/tissue type="lung"
/clone_lib="Lambda gt11"
163..1983
CDS

```

E38398		E38398	2866 bp	DNA	linear	PAT 31-JAN-2002
LOCUS		NF-kappa B activation inhibitory drug targeting TAK1 and method for identifying the same.				
DEFINITION		E38398				
ACCESSION		E38398.1	GI:18626978			
VERSION		JP 2000197500-A/4.				
KEYWORDS		unidentified				
SOURCE		unclassified.				
ORGANISM		1 (bases 1 to 2866)				
REFERENCE		Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.				
AUTHORS		NF-kappa B activation inhibitory drug targeting TAK1 and method for identifying the same				
TITLE		Patent: JP 2000197500-A 4 18-JUL-2000;				
JOURNAL		TANABE SEIYAKU CO LTD				
COMMENT		OS Unidentified				
		PN JP 2000197500-A/4				
		PD 18-JUL-2000				
		PF 04-FEB-1999 JP 1999026803				
		PI NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI HASEGAWA				
		PC C12Q1/48,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC C12Q1/02,				
		PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09, PC C12R1:91),				
		PC C12N5/00,C12N15/00,(C12N15/00,C12R1:91)				
		CC Strandedness: Double;				
		CC Topology: Linear;				
		FH Key Location/Qualifiers				
		FT source 1..2866 /organism='Unidentified'.				
FEATURES		Location/Qualifiers				
source		1..2866 /organism='unidentified'				
		/mol_type='genomic DNA'				
		/db_xref='taxon:32644'				
BASE COUNT		848 a 588 c 656 g 774 t				
ORIGIN						
		Query Match 100.0%; Score 684; DB 6; Length 2866;				
		Best Local Similarity 100.0%; Pred. No. 1.9e-200;				
		Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY		1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTGTAAAGCTTTATGAGGCC 60				
Db		388 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTGTAAAGCTTTATGAGGCC 447				
QY		61 TGCTTGAATCCAGTGTCCTTGTGATGGAATAATGCTGAAGGGGGCTCTTTATATAATGTG 120				
Db		448 TGCTTGAATCCAGTGTCCTTGTGATGGAATAATGCTGAAGGGGGCTCTTTATATAATGTG 507				
QY		121 CTGCATGTCGTGAACCAATTCGCAATTAATCTGTCGCCACGCAATGAGTTGGTGTATA 180				
Db		508 CTGCATGTCGTGAACCAATTCGCAATTAATTAATCTGTCGCCACGCAATGAGTTGGTGTATA 567				
QY		181 CAGTGTTCCTCAAGGAGTGCCTTATCTTCACAGCATGCAACCCAAAGCCGCTAATTCACAGG 240				
Db		568 CAGTGTTCCTCAAGGAGTGCCTTATCTTCACAGCATGCAACCCAAAGCCGCTAATTCACAGG 627				
QY		241 GACCTGAAACCACCAAACCTTACTGCTGTTGAGGGGGGACAGTTCCTAAAAAATTTGTGAT 300				
Db		628 GACCTGAAACCACCAAACCTTACTGCTGTTGAGGGGGGACAGTTCCTAAAAAATTTGTGAT 687				
QY		301 TTTGTCAGACCTGTGACATTCAGACACACATGACCAATTAACAAGGGGAGTGTGCTGTGG 360				
Db		688 TTTGTCAGACCTGTGACATTCAGACACACATGACCAATTAACAAGGGGAGTGTGCTGTGG 747				
QY		361 ATGGCACCTGAGTTTTTGAAGGTAGTAATTACAGTGAAMAAATGTGACGCTTCAGCTGG 420				
Db		748 ATGGCACCTGAGTTTTTGAAGGTAGTAATTACAGTGAAMAAATGTGACGCTTCAGCTGG 807				





This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 09:42:22 ; Search time 864.703 Seconds  
(without alignments)  
1911.298 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSPFTLQSTNTHQTSSSSS.....AEFYRLNSVDHGQSVVTAP 68

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09830144/runat\_03122003\_122509\_23332/app\_query.fasta\_1.654  
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 @CGN 1.1 3596 @runat\_03122003\_122509\_23332 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estmu:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mus:\*  
23: em\_gss\_man:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gsal:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	359	100.0	757	13	BX393376	BX393376	BX393376
2	359	100.0	974	10	BE898567	BE898567	BE898567
3	359	100.0	1021	13	BU152619	BU152619	BU152619
4	356	99.2	496	9	AW258218	AW258218	AW258218
5	356	99.2	578	13	BU609697	BU609697	BU609697
6	356	99.2	705	10	BE375985	BE375985	BE375985
7	355	98.9	782	12	BE253992	BE253992	BE253992
8	355	98.9	1385	13	BQ925425	BQ925425	BQ925425
9	353	98.3	997	10	BG115732	BG115732	BG115732
10	352	98.1	851	13	BX347763	BX347763	BX347763
11	351	97.8	334	10	BF710376	BF710376	BF710376
12	351	97.8	378	10	BG382061	BG382061	BG382061
13	351	97.8	527	12	BI400511	BI400511	BI400511
14	351	97.8	582	12	BI681846	BI681846	BI681846
15	351	97.8	599	12	BI401530	BI401530	BI401530
16	351	97.8	648	12	BI184403	BI184403	BI184403
17	351	97.8	815	12	BI181306	BI181306	BI181306
18	347	96.7	644	12	BM934363	BM934363	BM934363
19	345	96.1	777	12	BI181203	BI181203	BI181203
20	343	95.5	738	13	BU709271	BU709271	BU709271
21	339	94.4	930	13	BX347740	BX347740	BX347740
22	336	93.6	912	13	BX347764	BX347764	BX347764
23	327	91.1	919	10	BF163305	BF163305	BF163305
24	322	89.7	337	10	BF712308	BF712308	BF712308
25	311	86.6	325	9	AW658281	AW658281	AW658281
26	309	86.1	617	13	BU443120	BU443120	BU443120
27	309	86.1	570	10	BG710962	BG710962	BG710962
28	309	86.1	772	13	BU292972	BU292972	BU292972
29	309	86.1	941	13	BU427891	BU427891	BU427891
30	307	85.5	934	13	BX347741	BX347741	BX347741
31	288	80.2	677	13	BU426433	BU426433	BU426433
32	287	79.9	1107	10	BE902307	BE902307	BE902307
33	275	76.6	1002	14	BY709101	BY709101	BY709101
34	275	76.6	1576	11	AK009321	AK009321	AK009321
35	268	74.7	795	13	BU109063	BU109063	BU109063
36	267	74.4	450	9	AA674170	AA674170	AA674170
37	252	70.2	584	10	BE901639	BE901639	BE901639
38	251.5	70.1	610	13	BQ387712	BQ387712	BQ387712
39	247	68.8	695	13	BU256845	BU256845	BU256845
40	245.5	68.4	818	10	BE746542	BE746542	BE746542
41	245	68.2	654	13	BU417660	BU417660	BU417660
42	213	59.3	997	13	BX367448	BX367448	BX367448
43	210.5	58.6	1243	14	CD501051	CD501051	CD501051
44	187.5	52.2	931	14	CA469739	CA469739	CA469739
45	185	51.5	626	13	BX278260	BX278260	BX278260

# ALIGNMENTS

RESULT 1  
BX393376  
LOCUS  
DEFINITION  
BX393376 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CSDDC002Y002 5-PRIME, mRNA sequence.  
ACCESSION  
BX393376  
VERSION  
BX393376.1  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 757)

**AUTHORS** Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
**TITLE** Full-length cDNA libraries and normalization  
**JOURNAL** Unpublished  
**COMMENT** Contact: Genoscope

Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 10651.f For  
 more information about this cluster, see

[http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS08AM001ZC05\\_AW029\\_1&cluster=10651.f](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS08AM001ZC05_AW029_1&cluster=10651.f).

Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS08AM001ZC05\_AW029\_1.  
 Location/Qualifiers

# FEATURES

source

```
1..757
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS08AM001ZC05"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/Note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
```

BASE COUNT 162 a 253 c 197 g 145 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 6,01e-32 Length: 757  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x BX393376 (1-757)

```
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 33 CAAAGCCGACCTTAACCTTCAGTCAGTCACCAACAGCAGCAGCAGCAGCAGCTCCAGC 92
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 93 TCTGACGGAGGCTCTTCCTCCGCTCCGCGCCGCTCCGCTCCGCTCCGCTCCGCTCCG 152
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 153 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 212
QY 61 GluGlnSerValValThrAlaPro 68
Db 213 GAGCAGAGCGTGGTGACAGCACC 236
```

RESULT 2  
 BE898567  
 LOCUS BE898567 974 bp mRNA linear EST 29-SEP-2000  
 DEFINITION BE898567 NIH\_MGC\_9 Homo sapiens cDNA clone IMAGE:3951886 5', mRNA sequence.

ACCESSION BE898567  
 VERSION BE898567.1 GI:10365176  
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 974)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: DCTD/DRP

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)

Plate: LLCW820 row: k column: 23

High quality sequence stop: 652.

Location/Qualifiers

# FEATURES

source

```
1..974
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3951886"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/Note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 209 a 311 c 278 g 176 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 8,39e-32 Length: 974  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x BE898567 (1-974)

```
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 433 CAAAGCCGACCTTAACCTTCAGTCAGTCACCAACAGCAGCAGCAGCAGCAGCTCCAGC 492
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 493 TCTGACGGAGGCTCTTCCTCCGCTCCGCGCCGCTCCGCTCCGCTCCGCTCCGCTCCG 552
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 553 CGTGTTCAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 612
QY 61 GluGlnSerValValThrAlaPro 68
Db 613 GAGCAGAGCGTGGTGACAGCACC 636
```

## RESULT 3

BU152619

LOCUS BU152619

DEFINITION BU152619 1021 bp mRNA linear EST 03-SEP-2002

AGENCOURT 8678094 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6381058

5', mRNA sequence.

ACCESSION BU152619

VERSION BU152619.1 GI:226666151

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1021)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: DCTB/DTP  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LLCM2570 row: k column: 11  
 High quality sequence start: 48  
 High quality sequence stop: 601.  
 Location/Qualifiers  
 1..1021  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6381058"  
 /tissue\_type="carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_40"  
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
 BASE COUNT 205 a 386 c 259 g 165 t 6 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 8.93e-32 Length: 1021  
 Score: 359.00 Matches: 68  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x BU152619 (1-1021)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 244 CAAGCCCGAGCTTAAGCCCTGAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 303  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40  
 Db 304 TCTGACGAGGCGCTCTCCGCTCCGCGCCGCCACCTCGCTCCGCGAGGACGGT 363  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 Db 364 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 423  
 QY 61 GluGlnSerValThrAlaPro 68  
 Db 424 GAGCAGAGCGTGTGACAGCACCG 447

RESULT 4  
 AW258218 496 bp mRNA linear EST 03-APR-2000  
 LOCUS uq31h07.y1 NCI\_CGAP Mams Mus musculus cDNA clone IMAGE:2811037 5,  
 similar to TR:Q15750 Q15750 TAKI BINDING PROTEIN. ;, mRNA sequence.  
 DEFINITION  
 AW258218  
 ACCESSION  
 VERSION  
 KEYWORDS  
 EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 496)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished

## COMMENT

Other\_ESTS: uq31h07.x1  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 439.  
 Location/Qualifiers

FEATURES  
 source

1..496  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:2811037"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="7 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Mam5"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"  
 BASE COUNT 113 a 158 c 140 g 85 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 7.71e-32 Length: 496  
 Score: 356.00 Matches: 67  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 98.53% Mismatches: 0  
 Query Match: 99.16% Indels: 0  
 DB: 9 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x AW258218 (1-496)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 Db 17 CAGAGCCCCACTGACCCCTGAGTCCACCAACAGCAGCAGCAGCTCCAGC 76  
 QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40  
 Db 77 TCTGACGAGGCGCTCTCCGCTCCAGCCGCTCCTCCTCCACCGAGGATGGC 136  
 QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 Db 137 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGC 196  
 QY 61 GluGlnSerValThrAlaPro 68  
 Db 197 GAGCAGAGCGTGTGACGACCGCCT 220

## RESULT 5

BU609697 578 bp mRNA linear EST 20-FEB-2003  
 LOCUS UI-M-DJ2-bv-x-p-14-0-UI.r1 NIH BMAP DJ2 Mus musculus cDNA clone  
 DEFINITION UI-M-DJ2-bv-x-p-14-0-UI 5', mRNA sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 578)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery



ACCESSION BI253992  
 VERSION BI253992.1 GI:14805965  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 782)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: Incyte Genomics, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM11278 row: k column: 07  
 High quality sequence stop: 657.  
 Location/Qualifiers  
 1..782  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5114622"  
 /tissue\_type="cervical carcinoma cell line"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_12"  
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.4 kb. Library prepared by Life  
 Technologies."

BASE COUNT 162 a 288 c 214 g 118 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.84e-31 Length: 782  
 Score: 355.00 Matches: 67  
 Percent Similarity: 98.53% Conservative: 0  
 Best Local Similarity: 98.53% Mismatches: 1  
 Query Match: 98.89% Indels: 0  
 DB: 12 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x BI253992 (1-782)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 |||||  
 Db 306 CAAAGCCGACCTTACCTCGAGTCACCAACACGACGACGACGACGACGACGACGACG 365  
 |||||

QY 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 |||||  
 Db 366 TCTGACGAGGCGCTCTCCGCTCCGCGCCCTCGCCTCGCTCCGCGCGGAGGCGGT 425  
 |||||

QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 |||||  
 Db 426 CGTGTGAGCCCTATGTGGACTTTGTGAGTTTACCGCTCTGGAGCGTGGACCATGCG 485  
 |||||

QY 61 GluGlnSerValThrAlaPro 68  
 |||||  
 Db 486 GAGCAGAGCGTGTGACGACCG 509  
 |||||

RESULT 8  
 BQ25425  
 LOCUS  
 DEFINITION BQ25425 1385 bp mRNA linear EST 20-AUG-2002  
 AGENCOURT 8801162 NIH\_MGC\_129 Mus musculus cDNA clone IMAGE:6310974  
 5', mRNA sequence.  
 ACCESSION BQ25425  
 VERSION BQ25425.1 GI:22340456  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1385)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Susan L. Sullivan, PhD.  
 cDNA Library Preparation: ResGen, Invitrogen Corp  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM13734 row: c column: 07  
 High quality sequence start: 58  
 High quality sequence stop: 404.  
 Location/Qualifiers  
 1..1385  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6310974"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_129"  
 /note="Organ: olfactory epithelium; Vector:  
 pCMV-SPORT6.1.cdb; Site 1: EcoRV; Site 2: NotI; Cloned  
 unidirectionally. Primer: Oligo dT. Average insert size  
 2.2 kb. Constructed by ResGen, Invitrogen Corp. Note: this  
 is a NIH\_MGC Library."

BASE COUNT 257 a 536 c 337 g 253 t 2 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3.93e-31 Length: 1385  
 Score: 355.00 Matches: 67  
 Percent Similarity: 98.53% Conservative: 0  
 Best Local Similarity: 98.53% Mismatches: 1  
 Query Match: 98.89% Indels: 0  
 DB: 13 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x BQ25425 (1-1385)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
 |||||  
 Db 247 CAAAGCCGACCTTACCTCGAGTCACCAACACGACGACGACGACGACGACGACGACG 306  
 |||||

QY 21 SerAspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
 |||||  
 Db 307 TCTGACGAGGCGCTCTCCGCTCCGCGCCCTCGCCTCGCTCCGCGCGGAGGCGGT 366  
 |||||

QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60  
 |||||  
 Db 367 CGTGTGAGCCCTATGTGGACTTTGTGAGTTTACCGCTCTGGAGCGTGGACCATGCG 426  
 |||||

QY 61 GluGlnSerValThrAlaPro 68  
 |||||  
 Db 427 GAGCAGAGCGTGTGACGACCG 450  
 |||||

RESULT 9  
 BQ115732  
 LOCUS  
 DEFINITION BQ115732 997 bp mRNA linear EST 30-JAN-2001  
 602317018F1 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:4417108 5',  
 mRNA sequence.  
 ACCESSION BQ115732  
 VERSION BQ115732.1 GI:12609238  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.





SOURCE  
ORGANISM  
Sus scrofa (pig)  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE  
AUTHORS  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE  
Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477  
PUBMED  
889548  
COMMENT  
Contact: Tuggle CK  
Molecular Genetics Laboratory, Department of Animal Science  
Iowa State University  
201 Kildee Hall, Ames, IA 50011-3150, USA  
Tel: 5152944252  
Fax: 5152942401  
Email: cktuggle@iastate.edu  
Oligo-dT track not found. Not I site shown in beginning of sequence  
is likely internal to the message. cDNA Library Preparation: M.B.  
Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab,  
University of Iowa Clone distribution: clones will be available  
through Research Genetics (www.resgen.com) The following repetitive  
elements were found in this cDNA sequence: 62-112,  
>GC rich/Low complexity  
Seq primer: M13 Forward  
POLYA-No.

FEATURES  
source  
1..527  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/strain="crossbreed"  
/db\_xref="taxon:9823"  
/clone="MI-P-A11-nge-a-03-0-UI"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="MI-P-A11"  
/notes="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: EcoRI; The MI-P-A11  
library is normalized library derived from the MI-P-A10  
library, ultimately derived from placenta tissue. For a  
detailed description of the library from which this clone  
was derived, please visit our web site at  
http://pigest.genome.iastate.edu/. The procedure used to  
create this library has been previously described (Bonaldo  
, Lennon and Soares, Genome Research 6: 791-806, 1996)  
TAG SEQ=None found"  
BASE COUNT 67 a 166 c 216 g 78 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.22e-31 Length: 527  
Score: 351.00 Matches: 66  
Percent Similarity: 97.06% Conservative: 0  
Best Local Similarity: 97.06% Mismatches: 2  
Query Match: 97.77% Indels: 0  
DB: 12 Gaps: 0  
US-09-830-144-4\_COPY\_437\_504 (1-68) x BI400511 (1-527)  
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 318 CAGAGCCGCGACCTGACCTGACCTGCCAGTCCACCAACACCCAGAGAGAGAGTCCAGC 259  
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 258 TCCGACGGGGCCCTCTTCGCTCCCGGCCACCCACCTGCTCCCGCCCGGGGAGATGGC 199  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
Db 198 CGCGTGAGGCCCTACGTGGACTTTCGCGAGGTCTACCGGCTGTGGAGCGTGGACCGGC 139  
QY 61 GluGlnSerValValThrAlaPro 68

SOURCE  
ORGANISM  
Sus scrofa (pig)  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE  
AUTHORS  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE  
Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL  
Genome Res. 6 (9), 791-806 (1996)  
MEDLINE  
97044477  
PUBMED  
889548  
COMMENT  
Contact: Tuggle CK  
Molecular Genetics Laboratory, Department of Animal Science  
Iowa State University  
201 Kildee Hall, Ames, IA 50011-3150, USA  
Tel: 5152944252  
Fax: 5152942401  
Email: cktuggle@iastate.edu  
Oligo-dT track not found. Not I site shown in beginning of sequence  
is likely internal to the message. cDNA Library Preparation: M.B.  
Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab,  
University of Iowa Clone distribution: clones will be available  
through Research Genetics (www.resgen.com) The following repetitive  
elements were found in this cDNA sequence: 62-112,  
>GC rich/Low complexity  
Seq primer: M13 Forward  
POLYA-No.

FEATURES  
source  
1..527  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/strain="crossbreed"  
/db\_xref="taxon:9823"  
/clone="MI-P-A11-nge-a-03-0-UI"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="MI-P-A11"  
/notes="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: EcoRI; The MI-P-A11  
library is normalized library derived from the MI-P-A10  
library, ultimately derived from placenta tissue. For a  
detailed description of the library from which this clone  
was derived, please visit our web site at  
http://pigest.genome.iastate.edu/. The procedure used to  
create this library has been previously described (Bonaldo  
, Lennon and Soares, Genome Research 6: 791-806, 1996)  
TAG SEQ=None found"  
BASE COUNT 67 a 166 c 216 g 78 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.22e-31 Length: 527  
Score: 351.00 Matches: 66  
Percent Similarity: 97.06% Conservative: 0  
Best Local Similarity: 97.06% Mismatches: 2  
Query Match: 97.77% Indels: 0  
DB: 12 Gaps: 0  
US-09-830-144-4\_COPY\_437\_504 (1-68) x BI400511 (1-527)  
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 318 CAGAGCCGCGACCTGACCTGACCTGCCAGTCCACCAACACCCAGAGAGAGTCCAGC 259  
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 258 TCCGACGGGGCCCTCTTCGCTCCCGGCCACCCACCTGCTCCCGCCCGGGGAGATGGC 199  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60  
Db 198 CGCGTGAGGCCCTACGTGGACTTTCGCGAGGTCTACCGGCTGTGGAGCGTGGACCGGC 139  
QY 61 GluGlnSerValValThrAlaPro 68

Db 138 GAGCAGAGTGTGGCAGCGCGCGC 115

RESULT 14  
BI681846  
LOCUS  
DEFINITION  
461334 MARC IBOV Bos taurus cDNA 5', mRNA sequence.  
ACCESSION  
BI681846  
VERSION  
BI681846.1 GI:15634771  
KEYWORDS  
EST.  
SOURCE  
Bos taurus (cow)  
ORGANISM

REFERENCE  
AUTHORS  
1 (bases 1 to 582)  
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,  
Casas, E., Wray, J.B., White, J., Cho, J., Fahrnenkrug, S.C., Bennett  
G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G.,  
Perteau, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. and  
Keele, J.W.  
TITLE  
Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
JOURNAL  
Genome Res. 11 (4), 626-630 (2001)  
MEDLINE  
21180013  
PUBMED  
11282978  
COMMENT  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCCAGTCACGACG  
Plate: 135 row: 0 column: 15  
Seq primer: ATTAGTGACACTATAG.  
FEATURES  
source  
1..582  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="MARC IBOV"  
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from lymph node, ovary,  
fat, hypothalamus, and pituitary."  
BASE COUNT 93 a 212 c 179 g 98 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.67e-31 Length: 582  
Score: 351.00 Matches: 67  
Percent Similarity: 98.53% Conservative: 0  
Best Local Similarity: 98.53% Mismatches: 1  
Query Match: 97.77% Indels: 0  
DB: 12 Gaps: 0  
US-09-830-144-4\_COPY\_437\_504 (1-68) x BI681846 (1-582)  
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20  
Db 80 CAGAGCCGCGACCTGACCTGACCTGCCAGTCCACCAACACCCAGAGAGAGTCCAGC 139  
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40  
Db 140 TCCGACGGGGCCCTCTTCGCTCCCGGCCACCCACCTGCTCCCGCCCGGGGAGATGGC 199  
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60



Db 200 CCGGTGGAGCCCTACGTGGACTTTGCCGAGTTCTACCGCTCTGGAGCGTGGACCATGGC 259

QY 61 GluInSerValThrAlaPro 68

Db 260 GAGCAGAGTGTGTGACGGCGCG 283

RESULT 15

BI401530/c

LOCUS

DEFINITION MI-P-CP0-nvn-g-07-0-UI.s1 MI-P-CP0 Sus scrofa cDNA clone

ACCESSION BI401530

VERSION BI401530.1 GI:15180591

KEYWORDS EST.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

REFERENCE 1 (bases 1 to 599)

AUTHORS Ronaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

PUBMED 8889548

COMMENT Contact: Tuggle CK

Molecular Genetics Laboratory, Department of Animal Science

Iowa State University

201 Kildee Hall, Ames, IA 50011-3150, USA

Tel: 5152944252

Fax: 5152942401

Email: cktuggle@iastate.edu

Oligo-dT track not found. Not 1 site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab, University of Iowa Clone Distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 60-110,

>GC rich#Low complexity

Seq primer: M13 Forward

POLYA=No.

## FEATURES

Location/Qualifiers

1..599

/organism="Sus scrofa"

/mol\_type="mRNA"

/db\_xref="taxon:9823"

/clone="MI-P-CP0-nvn-g-07-0-UI"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="MI-P-CP0"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: EcoRI; The MI-P-CP0 library is derived from uterus. For a detailed description of the library from which this clone was derived, please visit our web site at <http://pigest.genome.iastate.edu/>. The procedure used to create this library has been previously described (Ronaldo, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG\_SEQ=None found"

BASE COUNT 77 a 193 c 238 g 91 t

ORIGIN

## Alignment Scores:

Pred. No.: 3..81e-31 Length: 599

Score: 351.00 Matches: 66

Percent Similarity: 97.06% Conservative: 0

Best Local Similarity: 97.06% Mismatches: 2

Query Match: 97.77% Indels: 0

DB: 12 Gaps: 0

US-09-830-144-4\_COPY\_437\_504 (1-68) x BI401530 (1-599)

QY 1 GluSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20

Db 316 CAGAGCCCCGACCCCTGACCTGCAGTCCACCAACACCCACAGAGCAGAGCTCCAGC 257

QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40

Db 256 TCCGACGGGGGCTCTTTTCGCTCCCGGCCACCCACTCGCTCCCGCCCGGAGATGGC 197

QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60

Db 196 CCGGTGGAGCCCTACGTGGACTTCGGGAGTTCTACCGCTCTGGAGCGTGGACCATGGC 137

QY 61 GluInSerValThrAlaPro 68

Db 136 GAGCAGAGTGTGGCAGCGGCCG 113

Search completed: December 4, 2003, 12:42:52

Job time : 867.703 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:21:16 ; Search time 28.4865 seconds  
(without alignments)  
615.997 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSPFTLQSTNTHTQSSSSS.....AEFYRLWSVDHGEQSVWTAP 68

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	356	99.2	500	11 Q8R0D1	Q8rd1 mus musculus
2	356	99.2	502	11 Q8CF89	Q8cf89 mus musculus
3	275	76.6	52	11 Q9CV62	Q9cv62 mus musculus
4	260.5	72.6	498	13 Q73614	Q73614 xenopus lae
5	74.5	20.8	329	4 Q96S04	Q96s04 homo sapien
6	70	19.5	1034	5 Q81QW0	Q81qw0 drosophila
7	70	19.5	1162	5 Q9VW87	Q9vw87 drosophila
8	69.5	19.4	389	12 Q9DXA1	Q9dxal avian pneum
9	69.5	19.4	389	12 Q9DXA0	Q9dxal avian pneum
10	67.5	18.8	703	4 Q9H451	Q9h451 homo sapien
11	67.5	18.8	739	4 Q43584	Q43584 homo sapien
12	67.5	18.8	862	4 Q9BV75	Q9bv75 homo sapien
13	67.5	18.8	903	4 Q96E66	Q96f66 homo sapien
14	66.5	18.5	497	16 Q8P7D0	Q8p7d0 xanthomonas
15	66	18.4	591	16 Q9RKD3	Q9rkd3 streptomyce
16	65.5	18.2	634	11 Q925Q8	Q925q8 mus musculus

17	65	18.1	452	12 Q91TB9	Q91tb9 hepatitis c
18	65	18.1	452	12 Q91TC8	Q91tc8 hepatitis c
19	65	18.1	452	12 Q91TC1	Q91tc1 hepatitis c
20	64.5	18.0	673	3 Q9C2H6	Q9c2h6 neurospora
21	64	17.8	452	12 Q91TC6	Q91tc6 hepatitis c
22	64	17.8	5604	4 Q8WZ53	Q8wz53 homo sapien
23	63.5	17.7	780	13 Q9FUI4	Q9fui4 xenopus lae
24	63.5	17.7	3021	12 Q92933	Q92933 hepatitis c
25	63.5	17.7	22152	4 Q8WXI7	Q8wxi7 homo sapien
26	63	17.5	452	12 Q91TC4	Q91tc4 hepatitis c
27	63	17.5	452	12 Q91TD5	Q91td5 hepatitis c
28	63	17.5	452	12 Q91TB8	Q91tb8 hepatitis c
29	63	17.5	452	12 Q91TB3	Q91td3 hepatitis c
30	63	17.5	452	12 Q91TC3	Q91td3 hepatitis c
31	63	17.5	452	12 Q91TB6	Q91tb6 hepatitis c
32	63	17.5	452	12 Q91TB4	Q91tb4 hepatitis c
33	63	17.5	2112	5 Q9VEL9	Q9vel9 drosophila
34	63	17.5	3021	12 Q81258	Q81258 hepatitis c
35	63	17.5	3021	12 Q68870	Q68870 hepatitis c
36	62.5	17.4	262	5 Q8MYC6	Q8myc6 apis mellif
37	62.5	17.4	268	5 Q8MYC4	Q8myc4 apis mellif
38	62.5	17.4	300	5 Q8MYC8	Q8myc8 apis mellif
39	62.5	17.4	303	5 Q8MYC7	Q8myc7 apis mellif
40	62.5	17.4	358	4 Q96BE7	Q96be7 homo sapien
41	62.5	17.4	393	2 Q8KY29	Q8ky29 streptomyce
42	62.5	17.4	1656	10 Q9LGS1	Q9lgs1 oryza sativ
43	62.5	17.4	1910	5 Q9V768	Q9v768 drosophila
44	62	17.3	172	10 Q9FGW8	Q9fgw8 arabidopsis
45	62	17.3	266	16 Q8FRA6	Q8fra6 corynebacte

## ALIGNMENTS

### RESULT 1

Q8R0D1 PRELIMINARY; PRT; 500 AA.  
ID Q8R0D1  
AC Q8R0D1  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Similar to mitogen-activated protein kinase kinase 7  
DE interacting protein 1 (fragment).  
GN 2310012M03RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
CX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC027054; AAH27054.1; -  
DR MGD: MGI:1913763; 2310012M03RIK.  
DR InterPro; IPR001932; PP2C-like.  
DR Pfam; PF00481; PP2C; 1.  
DR SMART; SM00332; PP2Cc; 1.  
KW Kinase.  
FT NON TER 1  
SQ SEQUENCE 500 AA; 54413 MW; 97529D2E2AE4DA34 CRC64;  
Query Match 99.2%; Score 356; DB 11; Length 500;  
Best Local Similarity 98.5%; Pred. No. 2e-34;  
Matches 67; Conservative 1; Mismatches 0; Gaps 0;

QY 1 QSPFTLQSTNTHTQSSSSS...AEFYRLWSVDHGEQSVWTAP 60  
Db 433 QSPFTLQSTNTHTQSSSSS...AEFYRLWSVDHGEQSVWTAP 492

QY 61 EQSVWTAP 68

Db 493 EQSVWTAP 500

```
RA Hayashizaki Y.;
RT "functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009321; BAB26215.1; -.
DR MGD; MGI:1913763; 2310012M03Rik.
FT NON TER 1
SQ SEQUENCE 52 AA; 5712 MW; 813E29B1639920A6 CRC64;

Query Match 76.6%; Score 275; DB 11; Length 52;
Best Local Similarity 98.1%; Pred. No. 8.2e-26;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 SSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEYRLWSVDHGQSVVTAP 68
Db 1 SSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEYRLWSVDHGQSVVTAP 52
[1]

RESULT 4
073614 PRELIMINARY; PRT; 498 AA.
AC 073614;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAB1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98120593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAK1 and TAB1 in BMP signaling in early Xenopus
development.";
RL EMBO J. 17:1019-1028(1998).
DR EMBL; U92031; RAC14009.1; -.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2CC; 1.
DR NCBI_TaxID=8355;
SQ SEQUENCE 498 AA; 54450 MW; 654AB8DE2353524 CRC64;

Query Match 72.6%; Score 260.5; DB 13; Length 498;
Best Local Similarity 73.5%; Pred. No. 6e-23;
Matches 50; Conservative 8; Mismatches 1; Gaps 1;

QY 1 QSPTLTQSTNTHTQSSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEYRLWSVDHG 60
Db 430 QSPSATLQSTNTHTQSSSSSDGLFRSRPLPSLPQDGEDGRVEPYVDFYRLWNAHN 489
[1]

QY 61 EQ-SVWTA 67
Db 490 DPGTLTA 497
[1]

RESULT 5
Q96S04 PRELIMINARY; PRT; 329 AA.
AC Q96S04;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN GS78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21096910; PubMed=11157797;
```

```
RA Hayashizaki Y.;
RT "functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009321; BAB26215.1; -.
DR MGD; MGI:1913763; 2310012M03Rik.
FT NON TER 1
SQ SEQUENCE 52 AA; 5712 MW; 813E29B1639920A6 CRC64;

Query Match 76.6%; Score 275; DB 11; Length 52;
Best Local Similarity 98.1%; Pred. No. 8.2e-26;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 SSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEYRLWSVDHGQSVVTAP 68
Db 1 SSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEYRLWSVDHGQSVVTAP 52
[1]

RESULT 4
073614 PRELIMINARY; PRT; 498 AA.
AC 073614;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAB1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98120593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAK1 and TAB1 in BMP signaling in early Xenopus
development.";
RL EMBO J. 17:1019-1028(1998).
DR EMBL; U92031; RAC14009.1; -.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2CC; 1.
DR NCBI_TaxID=8355;
SQ SEQUENCE 498 AA; 54450 MW; 654AB8DE2353524 CRC64;

Query Match 72.6%; Score 260.5; DB 13; Length 498;
Best Local Similarity 73.5%; Pred. No. 6e-23;
Matches 50; Conservative 8; Mismatches 1; Gaps 1;

QY 1 QSPTLTQSTNTHTQSSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEYRLWSVDHG 60
Db 430 QSPSATLQSTNTHTQSSSSSDGLFRSRPLPSLPQDGEDGRVEPYVDFYRLWNAHN 489
[1]

QY 61 EQ-SVWTA 67
Db 490 DPGTLTA 497
[1]

RESULT 5
Q96S04 PRELIMINARY; PRT; 329 AA.
AC Q96S04;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN GS78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21096910; PubMed=11157797;
```

RA Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,  
RA Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,  
RA Higgs D.R.;  
RT "Sequence, structure and pathology of the fully annotated terminal 2  
RT Mb of the short arm of human chromosome 16";  
RL Hum. Mol. Genet. 10:339-352(2001).  
DR EMBL, AF006465; AAK61262.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 329 AA; 35799 MW; 890FE4B3D1C5976D CRC64;  
Query Match 20.8%; Score 74.5; DB 4; Length 329;  
Best Local Similarity 28.6%; Pred. No. 0.9;  
Matches 22; Conservative 8; Mismatches 28; Indels 19; Gaps 2;  
QY 2 SPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGDEG-----RVPEY 45  
Db 138 APTTRICRTRCHPQSSGELSDG--RNCRPHDASENHRPGSSPVLGYFTRICKVERN 194  
QY 46 VDFAEFYRLWSVDHGEQ 62  
Db 195 IPECEDFRTWLTGSGEK 211  
RESULT 6  
Q81QW0 PRELIMINARY; PRT; 1034 AA.  
AC Q81QW0;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DE CG11940-Ph.  
GN CG11940.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abail J.F., Agbayani A., An H.J., Andrews-Frankoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Brothier P.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreria S., Fleischmann W.,  
RA Fosler K., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,

RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,  
RA Banon J., An H., Baldwin D., Banon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of Drosophila melanogaster genome";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of Drosophila melanogaster genome";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AE003513; AA09520.1; -;  
SQ SEQUENCE 1034 AA; 111096 MW; 827491ADDE5E4661 CRC64;  
Query Match 19.5%; Score 70; DB 5; Length 1034;  
Best Local Similarity 32.8%; Pred. No. 12;  
Matches 22; Conservative 9; Mismatches 24; Indels 12; Gaps 2;  
QY 3 PTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGDEGRVPEYDFAEFYRLWSVDHGE 61  
Db 932 FVLFPQRSPSTLTSCHSSSSAGSAYQTYAPGPMPLPPR-----ADVARLSSLSNGS 880  
QY 62 QSVWTAP 68  
Db 881 SSEVTSP 887  
RESULT 7  
Q9VWB7 PRELIMINARY; PRT; 1162 AA.  
ID Q9VWB7  
AC Q9VWB7;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE CG11940 protein.  
GN CG11940.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.

```

RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.B.G.,
RA Abril J.F., Adayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J.P., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brattier P.,
RA Burtis K.C., Busam D.A., Butler H.C., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davidson L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
SC Science 287:2185-2195(2000).

CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
DR EMBL; AE003513; AAF49029.1; -.
DR FlyBase; FBgn0031079; CG11940.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000159; RA_domain.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00314; RA; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
SQ SEQUENCE 1162 AA; 124386 MW; CA4C252D1E9AD795 CRC64;

Query Match 19.5%; Score 70; DB 5; Length 1162;
Best Local Similarity 32.8%; Pred. No. 13;
Matches 22; Conservative 9; Mismatches 24; Indels 12; Gaps 2;

QY 3 PTLTQSTNTHTQSSSSDGLFRS-PPAHSPLPFGEDGRVE 61
Db 960 PVLQPSPTLTSCHSSSSAGSAGYQTVAPGPMLEPR-----ADVRLSSLSNGS 1008

QY 62 QSVVTAP 68
Db 1009 SSVTSP 1015

RESULT 8
Q9DXA1 PRELIMINARY; PRT; 389 AA.
ID Q9DXA1
AC Q9DXA1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Attachment membrane glycoprotein.
GN G.

Query Match 19.4%; Score 69.5; DB 12; Length 389;
Best Local Similarity 37.2%; Pred. No. 4.4;
Matches 16; Conservative 7; Mismatches 17; Indels 3; Gaps 1;

QY 1 QSPTLTQSTNTHTQSSSSDGLFRSRPPAHSPLPFGEDGRVE 43
Db 274 RSPTPKQETGRAPRTATTQSG---SSPPHSSPPGVGDANME 313

RESULT 9
Q9DXA0 PRELIMINARY; PRT; 389 AA.
ID Q9DXA0
AC Q9DXA0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Attachment membrane glycoprotein.
GN G.

OS Avian pneumovirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OC NCBI_TaxID=38525;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fr/85/1;
RX MEDLINE=20495153; PubMed=11038385;
RX Bayon-Auboyer M.H., Arnault C., Toquin D., Etteradossi N.;
RT "Nucleotide sequences of the F, L and G protein genes of two non-
RT A/non-B avian pneumoviruses (APV) reveal a novel APV subgroup.";
RL J. Gen. Virol. 81:2723-2733(2000).
DR EMBL; AJ289466; CAC13042.1; -.
DR EMBL; AJ251085; CAC13041.1; -.
SQ SEQUENCE 389 AA; 41838 MW; 755CE4DEEC9C3EB8 CRC64;

Query Match 19.4%; Score 69.5; DB 12; Length 389;
Best Local Similarity 37.2%; Pred. No. 4.4;
Matches 16; Conservative 7; Mismatches 17; Indels 3; Gaps 1;

QY 1 QSPTLTQSTNTHTQSSSSDGLFRSRPPAHSPLPFGEDGRVE 43
Db 274 RSPTPKQETGRAPRTATTQSG---SSPPHSSPPGVGDANME 313

RESULT 10
Q9H451 PRELIMINARY; PRT; 703 AA.
ID Q9H451
AC Q9H451;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DJ46801.1 (Atrophin 1 interacting protein 4 (AIP4)) (Fragment).
GN DJ46801.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]

```



DR InterPro; IPR002349; WW.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 4.  
DR PRINTS; PR00403; WWDOMAIN.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECT; 1.  
DR SMART; SM00456; WW; 4.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 4.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 2.  
KW Ligase.  
SQ SEQUENCE 862 AA; 98675 MW; A3D960E7F4DBF9D3 CRC64;  
  
Query Match 18.8%; Score 67.5; DB 4; Length 862;  
Best Local Similarity 31.1%; Pred. No. 19;  
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;  
  
QY 2 SPTLTQSTNTHQTSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48  
Db 242 SSTGSLPPTNTINTSEGATSLIPLTISGSGRPLNPVTQAPLPQGEQRVD----- 296  
  
QY 49 AEFYRLWSVDHGQ 62  
Db 297 -QHGRVYVDHVEK 309  
  
RESULT 13  
Q96F66 PRELIMINARY; PRT; 903 AA.  
ID Q96F66  
AC Q96F66;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Similar to itchy (Mouse homolog) E3 ubiquitin protein ligase.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Tissue=Kidney;  
RA Scrausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
DR EMBL; BC011571; AAH11571.1; -.  
DR Genew; HGNC:13890; ITCB.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR000569; HECT\_domain.  
DR InterPro; IPR002349; WW.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 4.  
DR PRINTS; PR00403; WWDOMAIN.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECT; 1.  
DR SMART; SM00456; WW; 4.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 4.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 2.  
SQ SEQUENCE 903 AA; 102802 MW; 6777A2043C7B67BC CRC64;  
  
Query Match 18.8%; Score 67.5; DB 4; Length 903;  
Best Local Similarity 31.1%; Pred. No. 20;  
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;  
  
QY 2 SPTLTQSTNTHQTSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48  
Db 283 SSTGSLPPTNTINTSEGATSLIPLTISGSGRPLNPVTQAPLPQGEQRVD----- 337  
  
QY 49 AEFYRLWSVDHGQ 62  
Db 338 -QHGRVYVDHVEK 350  
  
RESULT 14  
Q9F7D0 PRELIMINARY; PRT; 497 AA.  
ID Q9F7D0  
AC Q9F7D0;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Outer membrane protein.  
GN OPRM OR XCC2681.  
OS Xanthomonas campestris (pv. campestris).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
OC Xanthomonadaceae; Xanthomonas.  
OX NCBI\_TaxID=340;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 33913 / NCPPB 528;  
MEDLINE=22022145; PubMed=12024217;  
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,  
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,  
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Clapina L.P.,  
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,  
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,  
RA Formighieri E.F., Franco M.C., Greggio C.C., Lemos M.V.F.,  
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira E.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
RA Setubal J.C., Kitajima J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
RT host specificities."  
RL Nature 417:459-463(2002).  
DR EMBL; AK012380; AAM41953.1; -.  
DR InterPro; IPR003423; OEP.  
DR Pfam; PF02321; OEP; 2.  
KW Complete proteome.  
SQ SEQUENCE 497 AA; 53173 MW; D41D5A1419B44C60 CRC64;  
  
Query Match 18.5%; Score 66.5; DB 16; Length 497;  
Best Local Similarity 36.1%; Pred. No. 13;  
Matches 26; Conservative 8; Mismatches 27; Indels 11; Gaps 2;  
  
QY 3 PTLTQSTNTHQTSSSSSDGGLFRSR-----PAHSLPPGEDGRVEPYVDFAEFRLW 55  
Db 327 PSISL-----TATLGSSSSSLGSLFESGTRAWSFVQLTLPLFNAGRNALDMAKANEDI 382  
  
QY 56 SVDHGQSVVTA 67  
Db 383 EVARYEKSQTA 394  
  
RESULT 15  
Q9RKD3 PRELIMINARY; PRT; 591 AA.  
ID Q9RKD3  
AC Q9RKD3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Putative integral membrane protein.  
GN SCO3154 OR SCE87.05.  
OS Streptomyces coelicolor.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;



RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Seeger K.J., Harris D.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2) / M145;  
RX MEDLINE=21996410; PubMed=12000953;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,  
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
coelicolor A3(2).";  
RL Nature 417:141-147(2002).  
DR EMBL: AL939115; CAB59650.1; --  
DR InterPro; IPR003342; Glyco\_trans\_39.  
DR Pfam; PF02366; PMT; 1.  
KW Complete proteome.  
SQ SEQUENCE 591 AA; 66289 MW; DE7F1707C0071897 CRC64;

Query Match 18.4%; Score 66; DB 16; Length 591;  
Best Local Similarity 47.4%; Pred. NO. 19;  
Matches 18; Conservative 2; Mismatches 14; Indels 4; Gaps 1;

QY 27 RSRPAHSLPPGCDGRVPEYVDFAEFYRL----WSVDHG 60  
||| | : || | | | | | | | | | : |  
Db 255 RSLAALPYDEDDGRVPEYVDFAEFYRL----WSVDHG 60  
||| | : || | | | | | | | | | : |

Search completed: December 4, 2003, 09:33:04  
Job time : 30.4865 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:23:11 ; Search time 11.7162 Seconds  
(without alignments)  
558.155 Million cell updates/sec

Title: US-09-830-144-4\_COPY\_437\_504

Perfect score: 359

Sequence: 1 QSPTLTQSTNTHTQSSSS.....AEFYLWSVDHGQSVVTAP 68

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 segs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63.5	17.7	243	D84792	probable protein t
2	62	17.3	541	S51799	nucleoporin NUP57
3	62	17.3	679	H95036	glycosyl hydrolase
4	62	17.3	737	D97907	alpha-xylosidase (
5	62	17.3	4957	T03455	ALR protein - huma
6	62	17.3	5262	T03454	ALR protein - huma
7	61.5	17.1	238	T52505	hypothetical prote
8	61.5	17.1	737	I39547	S-protein secretio
9	61	17.0	267	S74415	hypothetical prote
10	61	17.0	665	B82506	proteinase II VCA0
11	61	17.0	909	T06246	aspartate kinase (
12	61	17.0	951	B88042	protein F56D12.6 [
13	61	17.0	1511	A53151	pleiotropic drug r
14	60	16.7	1260	S60896	agglutinin-like pr
15	60	16.7	1576	AE0249	probable hemolysin
16	59.5	16.6	1369	T32338	hypothetical prote
17	59	16.4	173	G86861	hypothetical prote
18	59	16.4	329	S61884	carbonate dehydrat
19	59	16.4	330	S61883	carbonate dehydrat
20	59	16.4	330	S61882	carbonate dehydrat
21	59	16.4	429	T38146	dihydrofolate redu
22	59	16.4	651	JC7705	death receptor-6 -
23	58.5	16.3	338	H72679	hypothetical prote
24	58.5	16.3	378	A84161	hypothetical prote
25	58.5	16.3	608	I53269	prolactin receptor
26	58.5	16.3	1093	T38533	AP17 protein - hum
27	58.5	16.3	1099	T18257	phospholipase C -
28	58.5	16.3	2897	B48666	cell proliferation
29	58.5	16.3	3256	A48666	cell proliferation

30	58	16.2	142	2	D69891	ynkE protein - Bac
31	58	16.2	277	2	D69158	sensory transducti
32	58	16.2	532	2	B35621	spore germination
33	57.5	16.0	172	2	T36107	probable serine/ar
34	57.5	16.0	232	1	A25108	homeotic protein H
35	57.5	16.0	366	2	A86392	hypothetical prote
36	57.5	16.0	394	2	H65010	hypothetical prote
37	57.5	16.0	396	2	A57090	CSA protein - huma
38	57.5	16.0	629	2	B83107	chemotactic transd
39	57.5	16.0	686	2	JC5708	villin-like protei
40	57	15.9	762	2	S67765	probable membrane
41	57	15.9	778	2	H84678	hypothetical prote
42	56.5	15.7	276	2	T06331	photosystem II 22K
43	56.5	15.7	283	2	H83860	panthothenate synth
44	56.5	15.7	343	2	E83673	sorbitol dehydroge
45	56.5	15.7	457	2	B64790	ycbB protein - Esc

#### ALIGNMENTS

##### RESULT 1

D84792

probable protein translocase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001

C/Accession: D84792

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: D84792

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-243 <SFO>

A/Cross-references: GB:AE002093; NID:g4056494; PIDN:AAC98060.1; GSPDB:GN00139

C/Genetics:

A/Gene: At2g37410

A/Map position: 2

Query Match 17.7%; Score 63.5; DB 2; Length 243;  
Best Local Similarity 39.0%; Pred. No. 7.9;  
Matches 16; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

QY 8 QSTNTHTQSSSSDGGFLFRSPAHSLPPCGEDGRVEPYVDF 48

Db 193 QNQTASSSSSSSWFGGLF-DKKKEVQPGSEKTEVLESF 232

##### RESULT 2

S51799

nucleoporin NUP57 - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein G6320; protein YGR119C

C:Species: Saccharomyces cerevisiae

C/Date: 27-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 21-Jul-2000

C/Accession: S51799; S51800; S64428; S64427; S55976; S72192

R/Schlauch, N.L.

submitted to the EMBL Data Library, August 1994

A/Reference number: S51799

A/Accession: S51799

A/Molecule type: DNA

A/Residues: 1-541 <SCH>

A/Cross-references: EMBL:X81155; NID:g671635; PID:g671636

R/Grandi, P.; Schlaich, N.; Tekotte, H.; Hurt, E.C.

EMBO J. 14, 76-87, 1995

A/Title: Functional interaction of Nic96p with a core nucleoporin complex consisting of

A/Reference number: S51800; MUID:95129554; PMID:7828598

A/Accession: S51800

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-446, 'RL', 449-541 <GRA>

A;Cross-references: EMBL:X81155  
R;Van Dyck, L.; Skala, J.; de Wergifosse, P.; Purnelle, B.; Talla, E.; Nawrocki, A.; Del  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64428  
A;Accession: S64428  
A;Molecule type: DNA  
A;Residues: 1-353 <VAN>  
A;Cross-references: EMBL:Z72904; MIPS:YGR119C  
A;Experimental source: strain S288C  
R;Hansen, M.; Albers, M.; Backes, U.; Coblenz, A.; Leuther, H.; Neu, R.; Schreer, A.; Sc  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64417  
A;Accession: S64427  
A;Molecule type: DNA  
A;Residues: 300-541 <HAN>  
A;Cross-references: EMBL:Z72904; MIPS:YGR119C  
A;Experimental source: strain S288C  
R;van Dyck, L.; Goffeau, A.  
submitted to the EMBL Data Library, December 1994  
A;Title: The sequence of a 23.4 kb segment on the right arm of chromosome VII from Sacch  
A;Description: Genes for an asn synthase, a GLFG-motif nucleoporin and a putative homeob  
e new ORFs, remnants of Ty and three tRNA genes.  
A;Reference number: S55976  
A;Accession: S55976  
A;Molecule type: DNA  
A;Residues: 1-353 <VAV>  
A;Cross-references: EMBL:X83099  
R;Hansen, M.; Albers, M.; Backes, U.; Coblenz, A.; Leuther, H.; Neu, R.; Schreer, A.; Sc  
Yeast 12, 1273-1277, 1996  
A;Title: The sequence of a 23.4 kb segment on the right arm of chromosome VII from Sacch  
A;Reference number: S72179; MUID:97061913; PMID:8905931  
A;Accession: S72192  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-541 <HAN>  
A;Cross-references: EMBL:Z72904; MID:g1323192; PIDN:CAA97129.1; PID:g1323193  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1996  
C;Genetics:  
A;Gene: SGD:NUP57  
A;Cross-references: SGD:S0003351; MIPS:YGR119C  
A;Map position: 7R

Query Match 17.3%; Score 62; DB 2; Length 541;  
Best Local Similarity 50.0%; Pred. No. 30;  
Matches 12; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 8 QSTNTHTQSSSSSDGGLFRSRPA 31  
| | | : | | : | | | | |  
Dd 60 QATWTFGSNQSSSTGGGLFGNKPA 83

RESULT 3  
H95036  
glycosyl hydrolase, family 31 SP0312 [imported] - Streptococcus pneumoniae (strain TIGR4)  
C;Species: Streptococcus pneumoniae  
C;Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C;Accession: H95036  
R;Tattellin, H.; Neilson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,  
n, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A;Reference number: A95000; MUID:21357209; PMID:11463916  
A;Accession: H95036  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-679 <KUR>  
A;Cross-references: GB:AE005672; PIDN:AAK74489.1; PID:g14971785; GSPDB:GN00164; TIGR:SPA  
A;Experimental source: strain TIGR4  
C;Genetics:  
A;Gene: SP0312

Query Match 17.3%; Score 62; DB 2; Length 679;



A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: E82506  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-665 <HEI>  
 A:Cross-references: GB:AE004349; GB:AE003853; NID:g9657434; PIDN:AAF95977.1; GSPDB:GN001  
 A:Experimental source: serogroup O1; strain N16961; biotype E1 Tor  
 C:Genetics:  
 A:Gene: VCA0063  
 A:Map position: 2

Query Match 17.0%; Score 61; DB 2; Length 665;  
 Best Local Similarity 31.8%; Pred. No. 49;  
 Matches 14; Conservative 10; Mismatches 12; Indels 8; Gaps 2;

QY 32 HSLPFGEDGRVFPYVDF-----EFYRL--WSVDHGEQSVVTA 67  
 Db 96 HIVKGAQGOIEPLNLSARAEPPDYVQLASWSLDRSVQSVALA 139

RESULT 11  
 T06246  
 A:Title: aspartate kinase (EC 2.7.2.4) / homoserine dehydrogenase (EC 1.1.1.3) precursor - soybean  
 N:Contains: aspartate kinase (EC 2.7.2.4); homoserine dehydrogenase (EC 1.1.1.3)  
 C:Species: Glycine max (soybean)  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 21-Jan-2000  
 C:Accession: T06246  
 R:Gebhardt, J.S.; Weisemann, J.M.; Matthews, B.F.  
 A:Description: Genes encoding the bifunctional aspartokinase-homoserine dehydrogenase fr  
 A:Reference number: Z15563  
 A:Accession: T06246  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-909 <GEB>  
 A:Cross-references: EMBL:AF049708; NID:g2970554; PIDN:AAC05983.1; PID:g2970556  
 A:Experimental source: cultivar Century  
 C:Genetics:  
 A:Gene: AK-HSDH  
 A:Introns: 47/2; 69/1; 194/1; 223/3; 325/1; 342/3; 414/2; 443/3; 481/3; 514/3; 579/3; 60  
 C:Function: <AK>  
 A:Description: catalyzes phosphorylation of aspartate  
 C:Function: <HDH>  
 A:Description: catalyzes reduction of aspartate beta-semialdehyde into homoserine  
 C:Superfamily: chrA bifunctional enzyme; aspartate kinase homology; homoserine dehydroge  
 C:Keywords: multifunctional enzyme; oxidoreductase; phosphotransferase  
 F:87-552/Domain: aspartate kinase homology <DKI>  
 F:555-813/Domain: homoserine dehydrogenase homology <HSD>

Query Match 17.0%; Score 61; DB 2; Length 909;  
 Best Local Similarity 41.3%; Pred. No. 71;  
 Matches 19; Conservative 5; Mismatches 18; Indels 4; Gaps 2;

QY 2 SPTLTQSTNTHTQSSSSDGLFRSPAHSLPPG---EDGRVPE 44  
 Db 15 SPTLTLLHSHDRLPHSQC-RFFFLSRPSSHSLKGLTLPGRGAP 59

RESULT 12  
 E88042  
 A:Title: protein F56D12.6 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: E88042  
 R:anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites Genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C\_ele  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: E88042  
 A>Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-951 <STO>  
 A:Cross-references: GB:chr\_II; PIDN:AB66120.1; PID:g2315677; GSPDB:GN00020; CESP:F56D12  
 C:Genetics:  
 A:Gene: F56D12.6  
 A:Map position: 2

Query Match 17.0%; Score 61; DB 2; Length 951;  
 Best Local Similarity 43.9%; Pred. No. 74;  
 Matches 18; Conservative 5; Mismatches 14; Indels 4; Gaps 2;

QY 11 NTHQTSSSSSDGGLFR-SRPAHSL---PFGEDGRVPEYVD 47  
 Db 420 DTHSTASSKSDDKMLNGSAPAHSLDAPIDEKPKNLPVVD 460

RESULT 13  
 A53151  
 A:Title: pleiotropic drug resistance protein PDR5 - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: drug resistance protein YDR1; protein O3542; protein YOR153w; sporide  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 07-Jul-1995 #sequence\_revision 07-Jul-1995 #text\_change 02-Feb-2001  
 C:Accession: A53151; A49730; B49730; S48224; S55359; S67041; S34702  
 R:Bissinger, P.H.; Kuchler, K.  
 J. Biol. Chem. 269, 4180-4186, 1994  
 A:Title: Molecular cloning and expression of the Saccharomyces cerevisiae STS1 gene prod  
 A:Reference number: A53151; MUID:94140838; PMID:8307980  
 A:Accession: A53151  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1511 <BIS>  
 A:Cross-references: GB:X74113; NID:g395258; PIDN:CAA52212.1; PID:g395259  
 R:Balzi, E.; Wang, M.; Leteorne, S.; Van Dycck, L.; Goffeau, A.  
 J. Biol. Chem. 269, 2206-2214, 1994  
 A:Title: PDR5, a novel yeast multidrug resistance conferring transporter controlled by  
 A:Reference number: A49730; MUID:94124579; PMID:8294477  
 A:Accession: A49730  
 A:Molecule type: DNA  
 A:Residues: 1-1511 <BAL>  
 A:Cross-references: GB:L11922; NID:g402500; PIDN:AAB53769.1; PID:g402501  
 A:Accession: B49730  
 A:Molecule type: protein  
 A:Residues: 2-12 <BAW>  
 R:Hirata, D.; Yano, K.; Miyahara, K.; Miyakawa, T.  
 Curr. Genet. 26, 285-294, 1994  
 A:Title: Saccharomyces cerevisiae YDR1, which encodes a member of the ATP-binding casse  
 A:Reference number: S48224; MUID:95188264; PMID:7882421  
 A:Accession: S48224  
 A>Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-1511 <HIR>  
 R:Miyakawa, T.  
 submitted to the EMBL Data Library, January 1994  
 A:Reference number: S55359  
 A:Accession: S55359  
 A:Molecule type: DNA  
 A:Residues: 'M' 62-170, 'L', 172-189, 'I', 191-213, 'T', 215-307, 'V', 309-338, 345-475, 'H', 477-6  
 A:Cross-references: EMBL:D26548; NID:g57724; PIDN:BA05547.1; PID:d1006091; PID:g13216  
 R:Bordonne, R.; Camasses, A.; Madania, A.; Martin, R.P.; Poch, O.; Tarassov, I.A.; Wins  
 submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S67032  
 A:Accession: S67041  
 A:Molecule type: DNA  
 A:Residues: 1-1511 <BOR>  
 A:Cross-references: EMBL:Z75061; NID:g1420382; PIDN:CAA93359.1; PID:e252040; PID:g142038  
 C:Genetics:  
 A:Gene: SGD:PDR5; STS1; YDR1  
 A:Cross-references: SGD:S0005679; MIPS:YOR153w  
 A:Map position: 15R  
 C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology  
 C:Keywords: ATP; glycoprotein; nucleotide binding; P-loop; transmembrane protein  
 F:176-386/Domain: ATP-binding cassette homology <ABC1>  
 F:520-536/Domain: transmembrane #status predicted <TM1>



**This Page (uspto)**

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 09:43:02 ; Search time 91.6622 Seconds  
(without alignments)  
1097.894 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303

Perfect score: 1252

Sequence: 1 VELQLSRVNHPIVVKLYGA.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Xgapop 10.0 , Xgapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09830144/runat\_03122003\_122510\_23346/app.query.fasta\_1.654  
-DB=Issued\_Patents\_NA -QMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 @CGN 1 1 76 @runat\_03122003\_122510\_23346 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	100.0	1788	4	US-09-529-279-14
2	1252	100.0	1788	4	US-10-158-895-14
3	1252	100.0	2656	2	US-08-685-625A-5
4	1252	100.0	2656	4	US-09-529-279-3
5	1252	100.0	2656	4	US-10-158-895-3
6	1247	99.5	2443	2	US-08-685-625A-1
7	371	29.6	1365	3	US-09-221-235-6
8	371	29.6	1365	3	US-09-221-928-6
9	371	29.6	1365	3	US-09-221-527-6
10	371	29.6	1365	3	US-09-221-236-6
11	371	29.6	1365	3	US-09-221-416-6
12	371	29.6	1365	3	US-09-221-245-6

13	371	29.6	1365	3	US-09-163-115-6	Sequence 6, Appli
14	371	29.6	1365	3	US-09-221-528-6	Sequence 6, Appli
15	371	29.6	1365	3	US-09-593-553-6	Sequence 6, Appli
16	371	29.6	1365	3	US-09-221-237-6	Sequence 6, Appli
17	371	29.6	2119	4	US-09-399-588-1	Sequence 1, Appli
18	371	29.6	2120	3	US-09-221-235-4	Sequence 4, Appli
19	371	29.6	2120	3	US-09-221-928-4	Sequence 4, Appli
20	371	29.6	2120	3	US-09-221-527-4	Sequence 4, Appli
21	371	29.6	2120	3	US-09-221-236-4	Sequence 4, Appli
22	371	29.6	2120	3	US-09-221-416-4	Sequence 4, Appli
23	371	29.6	2120	3	US-09-221-245-4	Sequence 4, Appli
24	371	29.6	2120	3	US-09-163-115-4	Sequence 4, Appli
25	371	29.6	2120	3	US-09-221-528-4	Sequence 4, Appli
26	371	29.6	2120	3	US-09-593-553-4	Sequence 4, Appli
27	371	29.6	2120	3	US-09-221-237-4	Sequence 4, Appli
28	353.5	28.2	3389	1	US-08-395-580-1	Sequence 1, Appli
29	353.5	28.2	3426	1	US-08-205-018-1	Sequence 1, Appli
30	338.5	27.0	2505	3	US-09-291-839-3	Sequence 3, Appli
31	338.5	27.0	2505	4	US-09-458-457-3	Sequence 3, Appli
32	338.5	27.0	3025	3	US-09-291-839-1	Sequence 1, Appli
33	338.5	27.0	3025	4	US-09-458-457-1	Sequence 1, Appli
34	335.5	26.8	2505	4	US-09-458-457-9	Sequence 9, Appli
35	335.5	26.8	3026	4	US-09-458-457-7	Sequence 7, Appli
36	323	25.8	2890	1	US-07-928-464-1	Sequence 1, Appli
37	323	25.8	2890	5	PCT-US93-07347-1	Sequence 1, Appli
38	323	25.8	3033	1	US-08-003-311B-1	Sequence 1, Appli
39	323	25.8	3033	1	US-08-261-432-1	Sequence 1, Appli
40	305.5	24.4	1774	4	US-09-312-283C-403	Sequence 403, App
41	305.5	24.4	2370	4	US-09-509-802-1	Sequence 1, Appli
42	305.5	24.4	3516	3	US-09-188-930-257	Sequence 257, App
43	305.5	24.4	3516	4	US-09-312-283C-257	Sequence 257, App
44	304.5	24.3	4508	5	PCT-US93-06251-34	Sequence 34, Appl
45	294	23.5	1888	3	US-09-188-930-66	Sequence 66, Appl

ALIGNMENTS

RESULT 1  
US-09-529-279-14  
; Sequence 14, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1788  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)..(1776)  
US-09-529-279-14

Alignment Scores:  
Pred. No.: 4e-140  
Score: 1252.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
Length: 1788  
Matches: 228  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x US-09-529-279-14 (1-1788)

		FEATURE:	
		; NAME/KEY: CDS	
		; LOCATION: (7)..(1776)	
		US-10-158-895-14	
		Alignment Scores:	
		Pred. No.: 4e-140	
		Score: 1252.00	
		Percent Similarity: 100.00%	
		Best Local Similarity: 100.00%	
		Query Match: 100.00%	
		Indels: 0	
		Gaps: 0	
		DB: 4	
		US-09-830-144-2_COPY_76_303 (1-228) x US-10-158-895-14 (1-1788)	
QY	1	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleVallysLeuTyrGlyAla 20	1788
DB	232	GTAGAGCTTCGCGAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGCC 291	228
QY	21	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40	0
DB	292	TGCTTGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTATATAATGTG 351	0
QY	41	LeuHisGlyAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrrp 120	0
DB	532	TTTGGTACAGCTGTGACATTCAGACACATGACCAATAACAAGGGAGTGTGCTTGG 591	0
QY	121	MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrrp 140	0
DB	592	ATGGCACCTTGAAGGTAGTAAATACAGTGAAATAATGTGACGCTTCAGCTGG 651	0
QY	141	GlyIleLeuLeuTrrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160	0
DB	652	GGTATTATCTTTGGGAAGTGATAACCGTCGGAACCCCTTTGATGAGATTGGTGGCCCA 711	0
QY	161	AlaPheArgIleMetTrrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180	0
DB	712	GCCTTCGAATCATGTGGGCTGTTCATATGTTACTCGACCCACCTGATAAAAAATTTA 771	0
QY	181	ProLysProIleGluSerLeuMetThrArgCysTrrpSerLysAspProSerGlnArgPro 200	0
DB	772	CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGGTCTAAAGATCTCTCCAGCCCT 831	0
QY	201	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220	0
DB	832	TCAATGAGGAATTTGAAAAATAATGATCACTTGTATGCGGTACTTTCCAGGAGCAGAT 891	0
QY	221	GluProLeuGlnTrrpProCysGln 228	0
DB	892	GAGCCATTACAGTATCTCTGTCTAG 915	0
		RESULT 3	
		US-08-685-625A-5	
		; Sequence 5, Application US/08685625A	
		; Patent No. 5945301	
		; GENERAL INFORMATION:	
		; APPLICANT: UENO, Naoto	
		; APPLICANT: MATSUMOTO, Kunihiko	
		; APPLICANT: IRIE, Kenji	
		; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL	
		; TITLE OF INVENTION: TRANSDUCTION SYSTEM	

QY	1	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleVallysLeuTyrGlyAla 20	1788
DB	232	GTAGAGCTTCGCGAGTTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGAGCC 291	228
QY	21	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40	0
DB	292	TGCTTGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTATATAATGTG 351	0
QY	41	LeuHisGlyAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrrp 120	0
DB	532	TTTGGTACAGCTGTGACATTCAGACACATGACCAATAACAAGGGAGTGTGCTTGG 591	0
QY	121	MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrrp 140	0
DB	592	ATGGCACCTTGAAGGTAGTAAATACAGTGAAATAATGTGACGCTTCAGCTGG 651	0
QY	141	GlyIleLeuLeuTrrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160	0
DB	652	GGTATTATCTTTGGGAAGTGATAACCGTCGGAACCCCTTTGATGAGATTGGTGGCCCA 711	0
QY	161	AlaPheArgIleMetTrrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180	0
DB	712	GCCTTCGAATCATGTGGGCTGTTCATATGTTACTCGACCCACCTGATAAAAAATTTA 771	0
QY	181	ProLysProIleGluSerLeuMetThrArgCysTrrpSerLysAspProSerGlnArgPro 200	0
DB	772	CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGGTCTAAAGATCTCTCCAGCCCT 831	0
QY	201	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220	0
DB	832	TCAATGAGGAATTTGAAAAATAATGATCACTTGTATGCGGTACTTTCCAGGAGCAGAT 891	0
QY	221	GluProLeuGlnTrrpProCysGln 228	0
DB	892	GAGCCATTACAGTATCTCTGTCTAG 915	0
		RESULT 2	
		US-10-158-895-14	
		; Sequence 14, Application US/10158895	
		; Patent No. 6551840	
		; GENERAL INFORMATION:	
		; APPLICANT: ONO, KOICHIRO	
		; APPLICANT: OHTOMO, TOSHIHIKO	
		; APPLICANT: TSUCHIYA, MASAYUKI	
		; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES	
		; FILE REFERENCE: 053466/0278	
		; CURRENT APPLICATION NUMBER: US/10/158,895	
		; CURRENT FILING DATE: 2002-06-03	
		; PRIOR APPLICATION NUMBER: US/09/529,279	
		; PRIOR FILING DATE: 2000-04-11	
		; PRIOR APPLICATION NUMBER: PCT/JP98/04796	
		; PRIOR FILING DATE: 1998-10-22	
		; PRIOR APPLICATION NUMBER: JP 9/290188	
		; PRIOR FILING DATE: 1997-10-22	
		; NUMBER OF SEQ ID NOS: 48	
		; SOFTWARE: Patent In Ver. 2.1	
		; SEQ ID NO 14	
		; LENGTH: 1788	
		; TYPE: DNA	
		; ORGANISM: Homo sapiens	

NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
 STREET: P.O. Box 1404  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: United States  
 ZIP: 22313-1404  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/685,625A  
 FILING DATE: 24-JUL-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 7-253549  
 FILING DATE: 29-SEP-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meuth, Donna M.  
 REGISTRATION NUMBER: 36,607  
 REFERENCE/DOCKET NUMBER: 001560-267  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6620  
 TELEFAX: (703) 836-2021  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2656 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 183..1922  
 US-08-685-625A-5

Alignment Scores:  
 Pred. No.: 7,26e-140 Length: 2656  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-830-144-2\_COPY\_76\_303 (1-228) x US-08-685-625A-5 (1-2656)  
 QY 1 ValGluLeuArgGlnLeuSerArgValAlaHisProAsnIleValLysLeuTyrGlyAla 20  
 Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTTATGAGCC 467  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40  
 Db 468 TGCTTGAATCCAGTGTCTTGTGATGAATATGCTGAAGGGGGCTTTATATATG 527  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
 Db 528 CTCATGGTGGCTGAACCATTCGCATATTATCTGCTGCCACCAATGAGTTGGTGTGA 587  
 QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80  
 Db 588 CAGTGTTCACAGAGTGGCTTATCTTCAGCATGCAACCAAGCGCTTAATTCACAGG 647  
 QY 81 AspLeuLysProLysLeuLeuValAlaGlyThrValLeuLysIleCysAsp 100  
 Db 648 GACCTGAACACCAACCAATTTACTGCTGGTTGAGGGGGACACTTCTAAAAATTTGTGAT 707  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySerAlaAlaTyr 120  
 Db 708 TTTGGTACAGCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTGTGG 767

QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
 Db 768 ATGCACCTCTGAAGTTTTTGAAGGTAGTAATACAGTGAATAATGACGTCTTCAGCTGG 827  
 QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160  
 Db 828 GGTATTATTCTTTGGGAAGTGATAACCGTCGGAAACCTTTGATGAGATTGGTGGCCCA 887  
 QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180  
 Db 888 GCTTTCCGAATCATGTGGCTGTTTCAATATGTTACTCGACCACTGATAAAAAATTTA 947  
 QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200  
 Db 948 CCTAAGCCCATTTGAGACCTGATGACTCGTGTGTGGTCTAAGATCTCTCCACGCGCT 1007  
 QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220  
 Db 1008 TCAATGGAGGAATTTGAAAAATAAGTCACTTGTGATGCGTACTTTCAGGAGCAGAT 1067  
 QY 221 GluProLeuGlnTyrProCysGln 228  
 Db 1068 GAGCCATTACAGTATCTCTGTCTGAG 1091

RESULT 4  
 US-09-529-279-3  
 ; Sequence 3, Application US/09529279  
 ; Patent No. 6451617  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ONO, KOICHIRO  
 ; APPLICANT: OHTOMO, TOSHIHIKO  
 ; APPLICANT: TSUCHIYA, MASAYUKI  
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
 ; FILE REFERENCE: 053466/0278  
 ; CURRENT APPLICATION NUMBER: US/09/529,279  
 ; CURRENT FILING DATE: 2000-04-11  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
 ; PRIOR FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: JP 9/290188  
 ; PRIOR FILING DATE: 1997-10-22  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 2656  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (183)..(1919)  
 US-09-529-279-3

Alignment Scores:  
 Pred. No.: 7,26e-140 Length: 2656  
 Score: 1252.00 Matches: 228  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0  
 US-09-830-144-2\_COPY\_76\_303 (1-228) x US-09-529-279-3 (1-2656)

QY 1 ValGluLeuArgGlnLeuSerArgValAlaHisProAsnIleValLysLeuTyrGlyAla 20  
 Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTTATGAGCC 467  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40  
 Db 468 TGCTTGAATCCAGTGTCTTGTGATGAATATGCTGAAGGGGGCTTTTATATATG 527  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
 Db 528 CTGCATGGTGGCTGAACCATTCGCATATTATCTGCTGCCACCAATGAGTTGGTGTGA 587

Qy	61	GlnCysSerGlnGlyValIaIaTyLeuHisSerMetGlnProLysAlaLeuLeuHisArg	80
Db	588	CAGTGTTCACAAGAGTGGCTTATCTTCACAGCATCAACCCCAAGCGTAATTCACAGG	647
Qy	81	AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp	100
Db	648	GACCTGAACACCACAAACTTACTGTGTTGACAGGGGGGACAGTCTCTAAAATTTGTGAT	707
Qy	101	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIaTrp	120
Db	708	TTTGGTACAGCCTGTGACATTACAGACACACATGACCAATAACAAGCGGAGTGTCTGTGG	767
Qy	121	MetAlaProGluValPheGluGlySerAsnTySerGluLysCysAspValPheSerTrp	140
Db	768	ATGGCACTGAAGTTTTGAAGGTAGTAATACAGTGAATAAAATGTGACGCTCTTCACGTGG	827
Qy	141	GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro	160
Db	828	GGTATTATTCTTTGGGAAGTGATAACGCGTCGAAACCCCTTTGATGAGATGTGCTGCCCA	887
Qy	161	AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu	180
Db	888	GCITTCGAATCANGTGGGCTGTTTCATPAATGGTACTCGACACACACATGATAAAAAATTTA	947
Qy	181	ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro	200
Db	948	CCTAAGGCCCATTCAGAGCCCTGATGACTCGTTGTGTTGGTCTTAAGATCCCTTCCCAAGCGCCT	1007
Qy	201	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyPheProGlyAlaAsp	220
Db	1008	TCAATGGAGGAATGTGAAATATNAGCTCATCTTGATGCGGTACTTTCAGGAGCGAGAT	1067
Qy	221	GluProLeuGlnTyProCysGln	228
Db	1068	GAGCCATTACAGTATCCTTGTTCAG	1091

## RESULT 5

```

US-10-158-895-3
; Sequence 3, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-10-158-895-3

```

Alignment Scores:		
Pred. No.:	7.26e-140	2656
Score:	1252.00	228
Percent Similarity:	100.00%	Matches:
Best Local Similarity:	100.00%	Conservative:
Query Match:	100.00%	Mismatches:
DB:	4	Indels:
		Gaps:

US-09-830-144-2\_COPY\_76\_303 (11-228) x US-10-158-895-3 (11-2656)

QY	1	ValGluLeuArgGlnLeuSerArgValAenHisProAsnIleVallysLeuTyrGlyAla	20
DB	408	GTAGAGCTTCGGCAGTTATCCCGTGTGAACCACTCTAATATTGTAAAGCTTTATGAGCC	467
QY	21	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal	40
DB	468	TGCTTGAATCCAGTGTCTTGTGATTCGGAATATGCTGAAGGGGGCTTTATATAATATGG	527
QY	41	LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu	60
DB	528	CTGCATGGTGCTGAACCATTCGCATATTATACCTGTGCCACGCAATGAGTTGGTGTTTA	587
QY	61	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg	80
DB	588	CAGTGTTCCCAAGGAGTGCTTATCTTCACAGATGCAACCCAAAGCGCTAATTCACAGG	647
QY	81	AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp	100
DB	648	GACCTGAAACCAACCAACTTACTGCTGGTTGCAGGGGGACAGTTCTTAAAAATTTCTGTGAT	707
QY	101	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr	120
DB	708	TTTGGTACAGCCGTGACATTCAGACACATACCAATAACAGGGGAGTGCTGTTGG	767
QY	121	MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp	140
DB	768	ATGGCACCTGAGTTTTTGNAGGTAGTAATTACAGTGMAAAATGTGACGCTTTCAGCTGG	827
QY	141	GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro	160
DB	828	GGTATTATTCTTTGGGAAGTGATAACCGCTCGGAAACCCCTTTGATGAGATTGTTGGCCCA	887
QY	161	AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu	180
DB	888	GCITTCGGAATCATGTGGCTGTTCATATGTTACTCGACCCACCATGATPAAAAAATTTA	947
QY	181	ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro	200
DB	948	CCTAAGCCCCATTGAGAGCCCTGATGACTGCTGTTTGGTCTAAAGATCCTTCCAGCGCCCT	1007
QY	201	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp	220
DB	1008	TCAATGAGGAATTTGGAATAATAGACTCACTTGATGCGGTACTTTCAGGAGCAGAT	1067
QY	221	GluProLeuGlnTyrProCysGln	228
DB	1068	GAGCCATTACAGTATCTCTGTGAG	1091

DEPT. 6

RESULT 6  
 US-08-685-625A-1  
 ; Sequence 1, Application US/08685625A  
 ; Patent No. 5945301  
 ; GENERAL INFORMATION:  
 ; APPLICANT: UENO, Naoto  
 ; APPLICANT: MATSUMOTO, Kunihiko  
 ; APPLICANT: IRIE, Kenji  
 ; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL  
 ; TITLE OF INVENTION: TRANSDUCTION SYSTEM  
 ; NUMBER OF SEQUENCES: 5  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
 ; STREET: P.O. Box 1404  
 ; CITY: Alexandria  
 ; STATE: Virginia  
 ; COUNTRY: United States  
 ; ZIP: 22313-1404  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/685,625A

FILING DATE: 24-JUL-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-253549

FILING DATE: 29-SEP-1995

ATTORNEY/AGENT INFORMATION:

NAME: Meuth, Donna M.

REGISTRATION NUMBER: 36,607

REFERENCE/DOCKET NUMBER: 001560-267

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-6621

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2443 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 157..1893

US-08-685-625A-1

## Alignment Scores:

Pred. No.:	2,54e-139	Length:	2443
Score:	1247.00	Matches:	227
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.60%	Indels:	0
DB:	2	Gaps:	0

US-09-830-144-2\_COPY\_76\_303 (1-228) x US-08-685-625A-1 (1-2443)

QY	1	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla	20
DB	382	GTGAGCTCCGGCAGTTGTGCGGTGTAACCATCTTAACATTGTCAAGTTGTACGGAGCC	441
QY	21	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal	40
DB	442	TGCTGAATCCAGTATCTTGTGATGGAATATGCAGAGGGGGCTCATTTGATAATG	501
QY	41	LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu	60
DB	502	CTGCATGGTCTGAACCATTCCTTACTACACTGCTGCTCATGCCATGAGCTGGTGT	561
QY	61	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg	80
DB	562	CAGTGTCCCAAGAGTGGCTTACCTGCACAGATGCAGCCCAAGCGCTGATTCACAGG	621
QY	81	AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp	100
DB	622	GACCTCAAGCCTCCAACTTGTCTGCTGGTTCAGGAGGAGCAGTCTCAAAAATCTG	681
QY	101	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr	120
DB	682	TTTGGTACAGCTTGTGACATFCCAAACACATACCAATAATAAAGGAGTGTCTGTG	741
QY	121	MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr	140
DB	742	ATGGCGCTGAAGTGTGTAAGGTAGCAATTACAGTGAAAGTGTGATGCTTCAGCTGG	801
QY	141	GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro	160
DB	802	GATATTATCTCTGGGAAGTGATAACACGCGGAAACCCCTTCGATGAGATCGGTGCC	861
QY	161	AlaPheArgIleMetTyrPalValHisAsnGlyThrArgProProLeuIleLysAsnLeu	180
DB	862	GCTTTTCAAGATCATGTGGGCTGTTCATAATGGCACTCGACCACTGATCAAAAATTT	921
QY	181	ProLysProIleLeuSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro	200

DB	922	CCTAAGCCCATGTAGAGCTTGTATGACAGCTGTGGTCTAAGGACCCATCTCAGCCCT	981
QY	201	SerMetGluGluIleValIysLeuMetThrHisLeuMetArgTyrPheProGlyAlaAsp	220
DB	982	TCAATGGAGAAATGTGAAATATGACTCATTGATGCGGTACTTCCAGGAGCGGAT	1041
QY	221	GluProLeuGlnTyrProCys	227
DB	1042	GAGCCATTACAGTATCTTGT	1062

## RESULT 7

US-09-221-235-6  
; Sequence 6, Application US/09221235  
; Patent No. 6043040  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,235  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-235-6

## Alignment Scores:

Pred. No.:	9,19e-35	Length:	1365
Score:	371.00	Matches:	81
Percent Similarity:	56.28%	Conservative:	40
Best Local Similarity:	37.67%	Mismatches:	80
Query Match:	29.63%	Indels:	14
DB:	3	Gaps:	8

US-09-830-144-2\_COPY\_76\_303 (1-228) x US-09-221-235-6 (1-1365)

QY	2	GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCys	21
DB	157	GAGGCAGAAATACTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGAGTAATT	216
QY	22	LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn	39
DB	217	CTTGAACCTCCCAACTATGCGATTGTACAGAAATATGCTTCTCTGGGATCACTCTATG	276
QY	40	ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer	57
DB	277	TACATTAAACAGTAAACAGAGTGAGGAGTG-----GATATGATCATCATTTAGACC	327
QY	58	TyrCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu	77
DB	328	TGGGCCACTGATGTAGCCAAAGGAATGCATTATTATACATATGAGGCTCTGTCAAGTG	387
QY	78	IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys	97
DB	388	ATTCAAGAGACCTCAAGTCAAGAAACGTTGTATTAGCTGCTGATGGA---GTACTGAAG	444
QY	98	IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----	115
DB	445	ATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACATACACACATGCTCTTGGTT	501
QY	116	GlySerAlaAlaTyrMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys	135
DB	502	GGAACTTTCCCATGGATGGTCCAGAAAGTTATCCAGAGTCTCCCTGTGTGCAAACTTGT	561
QY	136	AspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAsp	155

Db 502 GGAACCTTTCCCATGGATGGCTCCAGAGTATATCCAGAGTCTCCCTGTGTGCAGAACTTGT 561  
QY 136 AspValPheSerTrpGlylleleuTrpGluValIleThrArgLysProPheasp 155  
Db 562 GACACATATCTCTATGGTGTGGTCTCTGGGAGATGCTAAACAGGAGGTGCCCTTTAAA 621  
QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174  
Db 622 GGTITGGAAGGA-----TTACAGTAGCTTGGCTTGTAGTGGAAGGAGAGATTA 675  
QY 175 ProLeuLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194  
Db 676 ACCATTCCAAGCAGTGGCCCAAGAGTTTGTGTAACCTGTTACATCAGTGTGGGAAGCT 735  
QY 195 AspProSerGlnArgProSerMetGluLleValLysIleMet 209  
Db 736 GATGCCAAGAAACGGCCATCATTCAGCAAAATCAITTCATCTG 780  
RESULT 9  
US-09-221-527-6  
; Sequence 6, Application US/09221527  
; Patent No. 6146832  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; CURRENT APPLICATION NUMBER: US/09/221,527  
; FILE REFERENCE: MNI-050  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-527-6  
Alignment Scores:  
Pred. No.: 9,19e-35 Length: 1365  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 3 Gaps: 8  
US-09-830-144-2\_COPY\_76\_303 (1-228) x US-09-221-527-6 (1-1365)  
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyGlyAlaCys 21  
Db 157 GAGGCGAATAACTAGTCTCCTCAGTCACAGAAACATCATCCAGTTTATGAGTAAT 216  
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyzAlaGluGlyGlySerLeuTyAsn 39  
Db 217 CTTGAACCTCCCACTAGTGCATGTGCACAGATATGCTTCTCGGATCACTCTATGAT 276  
QY 40 ValLeuHisGly-----AlaGluProLeuProTyTrpThrAlaAlaHisAlaMetSer 57  
Db 277 TACATTAAACAGTAACAGAGTGAGGAGATG-----GATATGGATCATCATATGACC 327  
QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyzLeuHisSerMetGlnProLysAlaLeu 77  
Db 328 TGGGCCACTGATGTAGCCAAAGGATGCATTTATACATATGAGGCTCTGTCAAGGTG 387  
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97  
Db 388 ATTCACAGACCTCAAGTCAAGAAAGCTTGTATTAGTCTGCTATGGA---GTACTGAAG 444  
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115

Db 562 GACACATATCTCTATGGTGTGGTCTCTGGGAGATGCTTAACAGGAGGTGCCCTTTAAA 621  
QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174  
Db 622 GGTITGGAAGGA-----TTACAGTAGCTTGGCTTGTAGTGGAAGGAGAGATTA 675  
QY 175 ProLeuLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194  
Db 676 ACCATTCCAAGCAGTGGCCCAAGAGTTTGTGTAACCTGTTACATCAGTGTGGGAAGCT 735  
QY 195 AspProSerGlnArgProSerMetGluLleValLysIleMet 209  
Db 736 GATGCCAAGAAACGGCCATCATTCAGCAAAATCAITTCATCTG 780  
RESULT 8  
US-09-221-928-6  
; Sequence 6, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; FILE REFERENCE: MNI-050  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-928-6  
Alignment Scores:  
Pred. No.: 9,19e-35 Length: 1365  
Score: 371.00 Matches: 81  
Percent Similarity: 56.28% Conservative: 40  
Best Local Similarity: 37.67% Mismatches: 80  
Query Match: 29.63% Indels: 14  
DB: 3 Gaps: 8  
US-09-830-144-2\_COPY\_76\_303 (1-228) x US-09-221-928-6 (1-1365)  
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyGlyAlaCys 21  
Db 157 GAGGCGAATAACTAGTCTCCTCAGTCACAGAAACATCATCCAGTTTATGAGTAAT 216  
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyzAlaGluGlyGlySerLeuTyAsn 39  
Db 217 CTTGAACCTCCCACTAGTGCATGTGCACAGATATGCTTCTCGGATCACTCTATGAT 276  
QY 40 ValLeuHisGly-----AlaGluProLeuProTyTrpThrAlaAlaHisAlaMetSer 57  
Db 277 TACATTAAACAGTAACAGAGTGAGGAGATG-----GATATGGATCATCATATGACC 327  
QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyzLeuHisSerMetGlnProLysAlaLeu 77  
Db 328 TGGGCCACTGATGTAGCCAAAGGATGCATTTATACATATGAGGCTCTGTCAAGGTG 387  
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97  
Db 388 ATTCACAGACCTCAAGTCAAGAAAGCTTGTATTAGTCTGCTATGGA---GTACTGAAG 444  
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115  
Db 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATACCAACACATCATGCTCTGGT 501  
QY 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyTrpSerGluLysCys 135

```

Db      445 ATCTGTGACCTTGGT---GCCTCTCGGTTCCATAACCATCAACACACATGCTCCTGGTT 501
Qy      116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135
Db      502 GGAACCTTCCCATGGATGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCCAGAACTTGT 561
Qy      136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgAlaGlyProPheAsp 155
Db      562 GACACATATCTATGTTGGTTCTCTGGGAGATGCTTAACAGGAGGTCCCTTTTAAA 621
Qy      156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
Db      622 GGTTTGAAGGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACAGAGATT 675
Qy      175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
Db      676 ACCATTCCAAAGCAGTTCGCCCCAGAAAGTTTGTCTGAAGTGTATACATCAGTGTGGGAAGCT 735
Qy      195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db      736 GATGCCAAGAAACGCCCATCATTTCAAGCAATCATTTCAATCCTG 780

RESULT 10
US-09-221-236-6
; Sequence 6, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-236-6

Alignment Scores:
Pred. No.: 9.19e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 3 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-236-6 (1-1365)
Qy      2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db      157 GAGCGAGAAATACTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT 216
Qy      22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db      217 CTTGAACCTCCCAACTATGCGCATGTGCAGAAATATGCTTCTCGGATCAGTCTATGAT 276
Qy      40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
Db      277 TACATTAAACAGTAACAGAAAGTAGGAGATG-----GATATGATCACATTATGACC 327
Qy      58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77
Db      328 TGGGCCACTGATGATGCCAAAGGAATGCAATTATTACATATGGAGGCTCCTGTCAAGGTG 387
Qy      78 IleHisArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97

```

```

Db      388 ATTCACAGAGACCTCAAGTCAAGAAACGTTGTTATAGTCTGTATGGA---GTACTGAAG 444
Qy      98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115
Db      445 ATCTGTGACCTTGGT---GCCTCTCGGTTCCATAACCATCAACACACATGCTCCTGGTT 501
Qy      116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135
Db      502 GGAACCTTCCCATGGATGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCCAGAACTTGT 561
Qy      136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgAlaGlyProPheAsp 155
Db      562 GACACATATCTATGTTGGTTCTCTGGGAGATGCTTAACAGGAGGTCCCTTTTAAA 621
Qy      156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
Db      622 GGTTTGAAGGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACAGAGATT 675
Qy      175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
Db      676 ACCATTCCAAAGCAGTTCGCCCCAGAAAGTTTGTCTGAAGTGTATACATCAGTGTGGGAAGCT 735
Qy      195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db      736 GATGCCAAGAAACGCCCATCATTTCAAGCAATCATTTCAATCCTG 780

RESULT 11
US-09-221-416-6
; Sequence 6, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-416-6

Alignment Scores:
Pred. No.: 9.19e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 3 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-416-6 (1-1365)
Qy      2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db      157 GAGCGAGAAATACTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT 216
Qy      22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db      217 CTTGAACCTCCCAACTATGCGCATGTGCAGAAATATGCTTCTCGGATCAGTCTATGAT 276
Qy      40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
Db      277 TACATTAAACAGTAACAGAAAGTAGGAGATG-----GATATGATCACATTATGACC 327
Qy      58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77

```







Thu Dec 4 17:00:23 2003

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21  
 Db 157 GAGGCAGAAATACTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT 216  
 QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39  
 Db 217 CTTGAACCTCCCACTATGGCATTGTACAGAAATATGCTTCTCTGGATCACTCTATGAT 276  
 QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57  
 Db 277 TACATTACAGTAACAGAGAGTACAGAGATG-----GATATGATCACATTATGACC 327  
 QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77  
 Db 328 TGGGCCACTGATGTAGCCAAAGATGATGATTTTACATATGAGGCTCTCTCAAGGTG 387  
 QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97  
 Db 388 ATTACAGAGACTCAAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---GTACTGAAG 444  
 QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115  
 Db 445 AICTGTGACTTTGGT---GCCCTCGGTTCCATACCATACACACACATGTCCTTGGTT 501  
 QY 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135  
 Db 502 GGAACCTTCCCATGGATGGTCCAGAAAGTTATCCAGAGTCTCCCTGTGTGAGAACTTGT 561  
 QY 136 AspValPheSerTrpGlyIleIleLeuTyrGluValIleThrArgArgLysPropheAsp 155  
 Db 562 GACACATATTCTTATGGTGTGGTCTCTGGAGATGCTAACAGGGAGGTCCCTTTAA 621  
 QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174  
 Db 622 GGTGTGAAGGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAACAGAGAGATTA 675  
 QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194  
 Db 676 ACCATTCCAAGCAGTTGCCCGCAGAAAGTTTGTGAACTGTTACATCATCTGTGGGAAGCT 735  
 QY 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209  
 Db 736 GATGCCCAAGAAACGGCCATCATTCAGCAAAATCATTTCAATCCCTG 780

Search completed: December 4, 2003, 12:44:58  
 Job time : 97.6622 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 04:36:53 ; Search time 84.7297 Seconds  
(without alignments)  
3563.166 Million cell updates/sec

Title: US-09-830-144-1\_COPY\_408\_1091

Perfect score: 684

Sequence: 1 gtagagcttgcgcagttatc.....cattacagatcctgtgcag 684

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:\*
- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
  - 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
  - 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
  - 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
  - 5: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
  - 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	684	100.0	1788	US-09-529-279-14	Sequence 14, Appl
2	684	100.0	1788	US-10-158-895-14	Sequence 14, Appl
3	684	100.0	2656	US-08-685-625A-5	Sequence 5, Appl
4	684	100.0	2656	US-09-529-279-3	Sequence 3, Appl
5	684	100.0	2656	US-10-158-895-3	Sequence 3, Appl
6	585.4	85.6	2443	US-08-685-625A-1	Sequence 1, Appl
7	54.6	8.0	1365	US-09-221-235-6	Sequence 6, Appl
8	54.6	8.0	1365	US-09-221-928-6	Sequence 6, Appl
9	54.6	8.0	1365	US-09-221-527-6	Sequence 6, Appl
10	54.6	8.0	1365	US-09-221-236-6	Sequence 6, Appl
11	54.6	8.0	1365	US-09-221-416-6	Sequence 6, Appl
12	54.6	8.0	1365	US-09-221-245-6	Sequence 6, Appl
13	54.6	8.0	1365	US-09-163-115-6	Sequence 6, Appl
14	54.6	8.0	1365	US-09-221-528-6	Sequence 6, Appl
15	54.6	8.0	1365	US-09-593-553-6	Sequence 6, Appl
16	54.6	8.0	1365	US-09-221-237-6	Sequence 6, Appl
17	54.6	8.0	2119	US-09-399-588-1	Sequence 6, Appl
18	54.6	8.0	2120	US-09-221-235-4	Sequence 1, Appl
19	54.6	8.0	2120	US-09-221-928-4	Sequence 4, Appl
20	54.6	8.0	2120	US-09-221-527-4	Sequence 4, Appl
21	54.6	8.0	2120	US-09-221-236-4	Sequence 4, Appl
22	54.6	8.0	2120	US-09-221-416-4	Sequence 4, Appl
23	54.6	8.0	2120	US-09-221-245-4	Sequence 4, Appl
24	54.6	8.0	2120	US-09-163-115-4	Sequence 4, Appl
25	54.6	8.0	2120	US-09-221-528-4	Sequence 4, Appl
26	54.6	8.0	2120	US-09-593-553-4	Sequence 4, Appl
27	54.6	8.0	2120	US-09-221-237-4	Sequence 4, Appl

28	52	7.6	1251	2	US-09-211-930-2	Sequence 2, Appl
29	52	7.6	1251	3	US-09-340-993-2	Sequence 2, Appl
30	52	7.6	1251	4	US-09-468-442-2	Sequence 2, Appl
31	52	7.6	1353	2	US-09-211-930-8	Sequence 8, Appl
32	52	7.6	1353	3	US-09-340-993-8	Sequence 8, Appl
33	52	7.6	1353	4	US-09-468-442-8	Sequence 8, Appl
34	52	7.6	1542	4	US-09-345-473E-13	Sequence 13, Appl
35	52	7.6	3201	2	US-09-211-930-1	Sequence 1, Appl
36	52	7.6	3201	3	US-09-340-993-1	Sequence 1, Appl
37	52	7.6	3201	4	US-09-468-442-1	Sequence 1, Appl
38	49.4	7.2	1979	4	US-09-685-462-3	Sequence 3, Appl
39	49.4	7.2	2028	2	US-09-211-930-12	Sequence 12, Appl
40	49.4	7.2	2028	3	US-09-340-993-12	Sequence 12, Appl
41	49.4	7.2	2028	4	US-09-468-442-12	Sequence 12, Appl
42	45.2	6.6	2890	1	US-07-928-464-1	Sequence 1, Appl
43	45.2	6.6	2890	5	PCT-US93-07347-1	Sequence 1, Appl
44	45.2	6.6	3033	1	US-08-003-311B-1	Sequence 1, Appl
45	45.2	6.6	3033	1	US-08-261-432-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-529-279-14  
; Sequence 14, Application US/09529279  
; Patent No. 6451617

; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI

; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279

; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22

; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 1788

; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (7)...(1776)

US-09-529-279-14

Query Match 100.0%; Score 684; DB 4; Length 1788;  
Best Local Similarity 100.0%; Pred. No. 1.6e-218;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GTAGAGCTTGGCAGTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGAGGCC	60
Db	232	GTAGAGCTTGGCAGTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGAGGCC	291
QY	61	TGCTTGAATCCAGTGCTCTGTGATGGAATATGCTGAAGGGGGCTTTATATATG	120
Db	292	TGCTTGAATCCAGTGCTCTGTGATGGAATATGCTGAAGGGGGCTTTATATATG	351
QY	121	CTGCATGCTGTGAACCATTTGCCATATATATCTGCTGCCACGCAATGATGGTGT	180
Db	352	CTGCATGCTGTGAACCATTTGCCATATATATCTGCTGCCACGCAATGATGGTGT	411
QY	181	CAGTGTTCACAGAGTGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAG	240
Db	412	CAGTGTTCACAGAGTGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAG	471
QY	241	GACCTGAACACCAACTTACTGCTGTTGAGGGGACAGTCTTAAATTTGTGAT	300
Db	472	GACCTGAACACCAACTTACTGCTGTTGAGGGGACAGTCTTAAATTTGTGAT	531



## TELECOMMUNICATION INFORMATION:

```

; ; TELEPHONE: (703) 836-6620
; ;
; ; TELEFAX: (703) 836-2021
; ;
; ; INFORMATION FOR SEQ ID NO: 5
; ;
; ; SEQUENCE CHARACTERISTICS:
; ;
; ; LENGTH: 2656 base pairs
; ;
; ; TYPE: nucleic acid
; ;
; ; STRANDEDNESS: double
; ;
; ; TOPOLOGY: linear
; ;
; ; MOLECULE TYPE: cDNA
; ;
; ; FEATURES:
; ;
; ; NAME/KEY: CDS
; ;
; ; LOCATION: 183..1922
; ;
; ; US-08-685-625A-5

```

Query Match	100.0%;	Score 684;	DB 2;	Length 2656;
Best Local Similarity	100.0%;	Pred. No. 2e-218;		
Matches 684;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GTAGAGCTTCGGCAGTATATCCCGTGTGAACCATCTCTAATATGTGAAGCTTTTATGGAGCC	60	
Db	408	GTAGAGCTTCGGCAGTATATCCCGTGTGAACCATCTCTAATATGTGAAGCTTTTATGGAGCC	467	
QY	61	TGCTTGAATCCAGTGTGTCCTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATGTG	120	
Db	468	TGCTTGAATCCAGTGTGTCCTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATGTG	527	
QY	121	CTGCATGGTGCTGAACCATTTGCCATATTAFACTGTGCTGCCACGCAATGACTTGTGTTTA	180	
Db	528	CTGCATGGTGCTGAACCATTTGCCATATTAFACTGTGCTGCCACGCAATGACTTGTGTTTA	587	
QY	181	CAGTGTTCACAGGAGTGGCTTATCTTCAGAGATGCAACCCAAAGCGCTTAATTCACAGG	240	
Db	588	CAGTGTTCACAGGAGTGGCTTATCTTCAGAGATGCAACCCAAAGCGCTTAATTCACAGG	647	
QY	241	GACCTGAACACCAACCAACTTACTGCTGGTTGTCAGGGGGGACACTTCTTAAAAATTTGTGAT	300	
Db	648	GACCTGAACACCAACCAACTTACTGCTGGTTGTCAGGGGGGACACTTCTTAAAAATTTGTGAT	707	
QY	301	TTTGGTACAGCGCTGTGCATTCAGACACACATGACCAATAACAAGGGGAGTGTGCTGTGG	360	
Db	708	TTTGGTACAGCGCTGTGCATTCAGACACACATGACCAATAACAAGGGGAGTGTGCTGTGG	767	
QY	361	ATGGCACCTGAAGTTTTTGAAGGTAGTAATTAACAGTGAAAAATGTACGCTCTTCAGCTGG	420	
Db	768	ATGGCACCTGAAGTTTTTGAAGGTAGTAATTAACAGTGAAAAATGTACGCTCTTCAGCTGG	827	
QY	421	GGTATTATTCTTTGGGAAGTGATAACCGGTTCGMAACCCCTTTGATGAGATTTGGTGCCCA	480	
Db	828	GGTATTATTCTTTGGGAAGTGATAACCGGTTCGMAACCCCTTTGATGAGATTTGGTGCCCA	887	
QY	481	GCTTTCCGAATCATGTGGGCTGTTTCATATATGTTACTTCGACCACTCATGATAAAAAATTTA	540	
Db	888	GCTTTCCGAATCATGTGGGCTGTTTCATATATGTTACTTCGACCACTCATGATAAAAAATTTA	947	
QY	541	CCTAAGCCCATTCAGAGCCTGATGACTCGTTGTTGTTCTTAAAGATCTCTTCCAGCGCCT	600	
Db	948	CCTAAGCCCATTCAGAGCCTGATGACTCGTTGTTGTTCTTAAAGATCTCTTCCAGCGCCT	1007	
QY	601	TCAATGGAGGAAATTTGTGAATAATATGACTCACTTGTATGTCGGTACTTTCACGAGCAGAT	660	
Db	1008	TCAATGGAGGAAATTTGTGAATAATATGACTCACTTGTATGTCGGTACTTTCACGAGCAGAT	1067	
QY	661	GAGCCATTACGATATCCTTTGTGAC	684	
Db	1068	GAGCCATTACGATATCCTTTGTGAC	1091	

## RESULT, T 4

US-09-529-279-3 ; Sequence 3, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:

; APPLICANT: ONO, KOICHIRO  
 ; APPLICANT: OHTOMO, TOSHIHIKO  
 ; APPLICANT: TSUCHIYA, MASAYUKI  
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
 ; FILE REFERENCE: 053466/0278  
 ; CURRENT APPLICATION NUMBER: US/09/529,279  
 ; CURRENT FILING DATE: 2000-04-11  
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
 ; PRIOR FILING DATE: 1998-10-22  
 ; PRIOR APPLICATION NUMBER: JP 9/290188  
 ; PRIOR FILING DATE: 1997-10-22  
 ; NUMBER OF SEQ ID NOS: 48  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 2656  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (183)..(1919)  
 ; US-09-529-279-3

Query Match	100.0.0%;	Score 684;	DB 4;	Length 2656;
Best Local Similarity	100.0.0%;	Pred. No. 2e-218;		
Matches 684;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	GTAGAGCTTCGGCAGTATCCCGTGTGCAACCAATCCTAAATATGTGTAAGCTTTATGGAGCC	60	
Db	408	GTAGAGCTTCGGCAGTATATCCCGTGTGCAACCAATCCTAAATATGTGTAAGCTTTATGGAGCC	467	
QY	61	TGCTTGAATCCAGTGTGCTTGTGTGATGGAAATATGCTGAAGGGGCTCTTTATATATAAATGTG	120	
Db	468	TGCTTGAATCCAGTGTGCTTGTGTGATGGAAATATGCTGAAGGGGCTCTTTATATATAAATGTG	527	
QY	121	CTGCATGTGTCTGAACCAATTGCCATATATPACTGTCTGCCACGCAATGAGTTGCTGTTTA	180	
Db	528	CTGCATGTGTCTGAACCAATTGCCATATATPACTGTCTGCCACGCAATGAGTTGCTGTTTA	587	
QY	181	CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAAATTCACAGG	240	
Db	588	CAGTGTTCCTCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAAATTCACAGG	647	
QY	241	GACCTGAAACCCAAACTTACTGCTGGTGTGCAGGGGGACAGTCTTCAAAAATTGTGAT	300	
Db	648	GACCTGAAACCCAAACTTACTGCTGGTGTGCAGGGGGACAGTCTTCAAAAATTGTGAT	707	
QY	301	TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAAGGGGAGTGTCTGTGG	360	
Db	708	TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAAGGGGAGTGTCTGTGG	767	
QY	361	ATGGCACCTGAGTGTTTTGAAGGTAGTAATTCAGTGAATAATGTGACGTTCTCAGCTGG	420	
Db	768	ATGGCACCTGAGTGTTTTGAAGGTAGTAATTCAGTGAATAATGTGACGTTCTCAGCTGG	827	
QY	421	GGTATATCTTTTGGGAAGTGATAACGCGTCGGAAACCTTTTGATGAGATTGCTGGCCCA	480	
Db	828	GGTATATCTTTTGGGAAGTGATAACGCGTCGGAAACCTTTTGATGAGATTGCTGGCCCA	887	
QY	481	GCTTTCGGAATCATGTGGCTGTTTCATATATGTTACTCGACCCACTGATAAAAAATTTA	540	
Db	888	GCTTTCGGAATCATGTGGCTGTTTCATATATGTTACTCGACCCACTGATAAAAAATTTA	947	
QY	541	CCTAAGCCCATTGAGAGCCTGATGACTCGTTTGGTCTAAAGATCTCTTCCAGGCGCCCT	600	
Db	948	CCTAAGCCCATTGAGAGCCTGATGACTCGTTTGGTCTAAAGATCTCTTCCAGGCGCCCT	1007	
QY	601	TCAATGGAGGAAATGTGAAAATAATGACTCATTGATGCGGTACTTCCACGAGCAGAT	660	
Db	1008	TCAATGGAGGAAATGTGAAAATAATGACTCATTGATGCGGTACTTCCACGAGCAGAT	1067	
QY	661	GAGCATTACAGTATCCTTGTGCAG	684	
Db	1068	GAGCATTACAGTATCCTTGTGCAG	1091	

948 CCTAAGCCCAATTGAGAGCCTGATGACTCGTTTGTGGTCTAAAGATCCTTCCAGGCGCCT 1007  
601 TCATGGAGGAAATGTGAAATATGATGACTCATTGATGCGGTACTTTCCAGGAGCAGAT 660  
1008 TCATGGAGGAAATGTGAAATATGATGACTCATTGATGCGGTACTTTCCAGGAGCAGAT 1067  
661 GAGCCATTACAGTATCCTTTGTCAG 684  
1068 GAGCCATTACAGTATCCTTTGTCAG 1091  
NO  
RESULT 6  
US-08-685-625A-1  
; Sequence 1, Application US/08685625A  
; Patent No. 5945301  
; GENERAL INFORMATION:  
; APPLICANT: UENO, Naoto  
; APPLICANT: MATSUMOTO, Kunihiko  
; APPLICANT: IRIE, Kenji  
; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL  
; TITLE OF INVENTION: TRANSDUCTION SYSTEM  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22313-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/685,625A  
; FILING DATE: 24-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-253549  
; FILING DATE: 29-SEP-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meuth, Donna M.  
; REGISTRATION NUMBER: 36,607  
; REFERENCE/DOCKET NUMBER: 001560-267  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-8620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2443 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 157..1893  
; US-08-685-625A-1  
Query Match 85.6%; Score 585.4; DB 2; Length 2443;  
Best Local Similarity 91.1%; Pred. No. 1.8e-185;  
Matches 622; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGGAGCC 60  
DB 382 GTGAGCTCCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGGAGCC 441  
QY 61 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
DB 442 TGCCTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTATTGATAATGTG 501  
QY 121 CTGCATGGTGTGAACCAATTGCCATATTTATCTGCTGCCACCAATGAGTTGGTTTA 180

US-10-158-895-3  
; Sequence 3, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO: 3  
; LENGTH: 2656  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (183)..(1919)  
US-10-158-895-3  
Query Match 100.0%; Score 684; DB 4; Length 2656;  
Best Local Similarity 100.0%; Pred. No. 2e-218;  
Matches 684; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGGAGCC 60  
DB 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATTTGTAAGCTTTATGGAGCC 467  
QY 61 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 120  
DB 468 TCGTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 527  
QY 121 CTGCATGGTGTGAACCAATTGCCATATTTATCTGCTGCCACCAATGAGTTGGTTTA 180  
DB 528 CTGCATGGTGTGAACCAATTGCCATATTTATCTGCTGCCACCAATGAGTTGGTTTA 587  
QY 181 CAGTGTTCACAGAGTGGCTTATCTTCAGCATGCAACCCAAAGCGCTTAATTCACAGG 240  
DB 588 CAGTGTTCACAGAGTGGCTTATCTTCAGCATGCAACCCAAAGCGCTTAATTCACAGG 647  
QY 241 GACCTGAAACACCAACTTACTGCTGTTGCGAGGGGGGACAGTTCTAAATAATTTGAT 300  
DB 648 GACCTGAAACACCAACTTACTGCTGTTGCGAGGGGGGACAGTTCTAAATAATTTGAT 707  
QY 301 TTTGGTACAGCTGTGACATTCAGACACATGACCAATACAGGGGAGTGTCTGTG 360  
DB 708 TTTGGTACAGCTGTGACATTCAGACACATGACCAATACAGGGGAGTGTCTGTG 767  
QY 361 ATGGCAGCTGAAGTTTGTGAGGTAGTAAATACAGTGAATAATGAGCTTCTCAGCTGG 420  
DB 768 ATGGCAGCTGAAGTTTGTGAGGTAGTAAATACAGTGAATAATGAGCTTCTCAGCTGG 827  
QY 421 GGTATTATCTTTGGGAGTGATTAACCGCTCGGAAACCTTTGATGAGATTGTTGGCCCA 480  
DB 828 GGTATTATCTTTGGGAGTGATTAACCGCTCGGAAACCTTTGATGAGATTGTTGGCCCA 887  
QY 481 GCTTTCCGATCATGTGGGCTGTTCAATGCTGACCACTGACCACTGATATAAATTTA 540  
DB 888 GCTTTCCGATCATGTGGGCTGTTCAATGCTGACCACTGACCACTGATATAAATTTA 947  
QY 541 CCTAAGCCCAATTGAGAGCTGATGACTCGTTTGTGGTCTAAAGATCCTTCCAGGCGCT 600

Db	502	CTGCATGGTGTGAACCAATTGCCCTTACACTGCTGCTATGCCATGCGATGAGCTGTGTTTA	561
Qy	181	CAGTGTTCCTCAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG	240
Db	562	CAGTGTTCCTCAGGAGTGGCTTACCTGCACAGCATGACGCCCAAGCGCTGATTCACAGG	621
Qy	241	GACCTGAAACCAACCAACTTACTGCTGTTGAGGGGGAAGTTTCTAAAAATTTGTGAT	300
Db	622	GACCTCAAGCCTCCAACTTCTGCTGTTGCAGGAGGACAGTTCTAAAAATCTGCGAT	681
Qy	301	TTTGGTACAGCCTGTGCAATTGAGACACACATGACCAATACAGGGGAGTGGCTGTGG	360
Db	682	TTTGGTACAGCTTGTGATCCAAACACACATGACCAATAATAAGGGAGTGGCTGTGG	741
Qy	361	ATGGCACTCTGAAGTTTTGAAGGTAGTAATTACAGTGAATAATGTGACGCTCTCAGCTGG	420
Db	742	ATGGCGCTCTGAAGTGTTTGAAGGTAGCAATTTACAGTGAATAAGTGTGATGCTTCAGCTGG	801
Qy	421	GGTATTATTCTTTGGGAAGTGATAACCGTGCGAAAAACCCCTTTGATGAGATGGTGGGCCA	480
Db	802	GGTATTATCTCTGGGAAGTGATAACACGCGGAAACCTTCGATGAGATCGGTGGGCCA	861
Qy	481	GCCTTCGGAATCATGTGGGCTGTTTCATAATGGTACTCGACACCACTGATAAAAAATTA	540
Db	862	GCCTTCAGAAATCATGTGGCTGTTTCATAATGGCACTCGACCACTGATCAAAAAATTA	921
Qy	541	CCCTAAGCCCATTCAGAGCCTGATGATCTGTTGTTGGTCTAAGATCCTTCCACGCGCCT	600
Db	922	CCCTAAGCCCATTCAGAGCTTGATGACACGCTGTGGTCTAAGAGACCATCTCAGCGCCT	981
Qy	601	TCAATGGAGAAATTGTGAAAAATAATGACATCACTTGATCGGTACTTCCAGGAGCAGAT	660
Db	982	TCAATGGAGAAATTGTGAAAAATAATGACTCACTTGATCGGTACTTCCAGGAGCGGAT	1041
Qy	661	GAGCCATTACAGTATCCTTGTC	683
Db	1042	GAGCCATTACAGTATCCTTGTC	1064

```

RESULT 7
US-09-221-235-6
; Sequence 6, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
; US-09-221-235-6

```

	Query Match	8.0%;	Score 54.6;	DB 3;	Length 1365;
	Best Local Similarity	48.3%;	Pred. No. 3.8e-08;		
	Matches 153;	Conservative 0;	Mismatches 164;	Indels 0;	Gaps 0;
160	CACGCAATGAGTTGGTGTTTACAGTGTTC	CCCCAAGGAGTGGCTTATCTTCACAGCATGCA	219		
316	CACATTATGACCTGGGCCACTGATGCGCAAGGAATGCATATTATTACATATGGAGGCT	375			
220	CCCAAAGCGCTAAATTCACAGGACCTGAAACCAACCAACTTACTCTGTTGTCAGGGGGG	279			

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,527  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-527-6

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,416  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-416-6

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,236  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-236-6

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,527  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-527-6

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,416  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-416-6

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,236  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-236-6



Db 556 ACTTGTGACACATATTCCTATGGTGTCTCTGGGAGATGCTAACCAAGGAGGTCCCC 615  
QY 460 TTGTGATGAGATTGGGG 476  
Db 616 TTTAAAGGTTTGAAGG 632

## RESULT 12

US-09-221-245-6  
; Sequence 6, Application US/09221245  
; Patent No. 6180358  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,245  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-245-6

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
QY 160 CAGCATGAGTTGGTGTACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAA 219  
Db 316 CACATTATGACCTGGGCCACTGATGTAGCCAAAGGAATGCATTATTACATATGGAGCT 375  
QY 220 CCCAAAGCGCTAATTCACAGGACCTGAACACCACCAACTTACTGCTGGTGGAGGGGG 279  
Db 376 CCTGTCAAGGTGATTTCACAGAGACCTCAAGTCAAGAAACGTTGTATAGCTGCTGATGA 435  
QY 280 ACAGTTCTAAAAATTTGTGATTTTGTGTACAGCCTGTGACATTCAGACACACATGACCAAT 339  
Db 436 GTACTGAAGATCTGTGACTTTGGTGGCTCTCGGTTCCATACCATACACACATGTCC 495  
QY 340 AACAGGGGAGTGGCTGTGGATGGCACCTGAAGTTTTTGAAGGTAGTAATTACAGTGAA 399  
Db 496 TTGGTTGGAACCTTCCCATGGATGGCTCCAGAGTTATCCAGAGTCTCCCTGTGTGAGAA 555  
QY 400 AAATGTGACGCTCTTCAGCTGGGTATTTCTTTGGGAAGTATGATACGGCTCGGAACCC 459  
Db 556 ACTGTGACACATATTCCTATGGTGTGTTCTCTGGGAGATGCTAACAGGGAGGTCCCC 615  
QY 460 TTGTGATGAGATTGGTG 476  
Db 616 TTTAAAGGTTTGAAGG 632

## RESULT 13

US-09-163-115-6  
; Sequence 6, Application US/09163115A  
; Patent No. 6181962  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/163,115A  
; CURRENT FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6

; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-163-115-6

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
QY 160 CAGCATGAGTTGGTGTACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAA 219  
Db 316 CACATTATGACCTGGGCCACTGATGTAGCCAAAGGAATGCATTATTACATATGGAGCT 375  
QY 220 CCCAAAGCGCTAATTCACAGGACCTGAACACCACCAACTTACTGCTGGTGGAGGGGG 279  
Db 376 CCTGTCAAGGTGATTTCACAGAGACCTCAAGTCAAGAAACGTTGTATAGCTGCTGATGA 435  
QY 280 ACAGTTCTAAAAATTTGTGATTTTGTGTACAGCCTGTGACATTCAGACACACATGACCAAT 339  
Db 436 GTACTGAAGATCTGTGACTTTGGTGGCTCTCGGTTCCATACCATACACACATGTCC 495  
QY 340 AACAGGGGAGTGGCTGTGGATGGCACCTGAAGTTTTTGAAGGTAGTAATTACAGTGAA 399  
Db 496 TTGGTTGGAACCTTCCCATGGATGGCTCCAGAGTTATCCAGAGTCTCCCTGTGTGAGAA 555  
QY 400 AAATGTGACGCTCTTCAGCTGGGTATTTCTTTGGGAAGTATGATACGGCTCGGAACCC 459  
Db 556 ACTGTGACACATATTCCTATGGTGTGTTCTCTGGGAGATGCTAACAGGGAGGTCCCC 615  
QY 460 TTGTGATGAGATTGGTG 476  
Db 616 TTTAAAGGTTTGAAGG 632

## RESULT 14

US-09-221-528-6  
; Sequence 6, Application US/09221528  
; Patent No. 6190874  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,528  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 1365  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1365)  
US-09-221-528-6

Query Match 8.0%; Score 54.6; DB 3; Length 1365;  
Best Local Similarity 48.3%; Pred. No. 3.8e-08;  
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;  
QY 160 CAGCATGAGTTGGTGTACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAA 219  
Db 316 CACATTATGACCTGGGCCACTGATGTAGCCAAAGGAATGCATTATTACATATGGAGCT 375  
QY 220 CCCAAAGCGCTAATTCACAGGACCTGAACACCACCAACTTACTGCTGGTGGAGGGGG 279  
Db 376 CCTGTCAAGGTGATTTCACAGAGACCTCAAGTCAAGAAACGTTGTATAGCTGCTGATGA 435  
QY 280 ACAGTTCTAAAAATTTGTGATTTTGTGTACAGCCTGTGACATTCAGACACACATGACCAAT 339

Thu Dec 4 17:00:20 2003

```

Db      436 GTACTGAAGATCTGTGACTTTGGTCCCTCTCGGTTCCATTAACCATACACACATGTCC 495
QY      340 AACAGGGGAGTGTGCTGTGATGGACCTGAAGTTTTTGAAGGTAGTAAATTACAGTGAA 399
Db      496 TTGGTTGGAACCTTCCCAATGATGGCTCCAGAAGTTATCCAGAGTCTCCCTGTGTGAGAA 555
QY      400 AAATGTGACGCTTTCAGTGGGATTAATTTCTTTGGGAAGTGATAACGGTCGGAAACCC 459
Db      556 ACTTGTGACACATATTCTATGGTGTGTTCTCTGGGAGATGCTAAACAAGGAGGTCCCC 615
QY      460 TTTGATGAGATTGGTGG 476
Db      616 TTTAAAGGTTTGAAGG 632

```

```

RESULT 15
US-09-593-553-6
; Sequence 6, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-593-553-6

```

```

Query Match      8.0%; Score 54.6; DB 3; Length 1365;
Best Local Similarity 48.3%; Pred. No. 3.8e-08;
Matches 153; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY      160 CACGCAATGAGTTGGTGTACAGTGTCCCAAGGAGTGGCTTATCTTCACAGCATGCAA 219
Db      316 CACATTATGACCTGGGCCACTGATGTAGCCAAAGGATGCATTATTTACATATGGAGCT 375
QY      220 CCCAAGCGCTAATTCACAGGACCTGAACCCACCAACTTACTGCTGGTTGCAGGGGG 279
Db      376 CCTGTCAGGTGATTCACAGAGACCTCAAGTCAAGAACGTTGTATAGCTGTGATGGA 435
QY      280 ACAGTCTCTAAAAATTTGTGATTTTGTGACAGCCTGTGACATTCAGACACACATGACCAAT 339
Db      436 GTACTGAAGATCTGTGACTTTGGTGGCTCTCGGTTCCATAACCATACACACATGTCC 495
QY      340 AACAGGGGAGTGTGCTGTGATGGACCTGAAGTTTTTGAAGGTAGTAAATTACAGTGAA 399
Db      496 TTGGTTGGAACCTTCCCAATGATGGCTCCAGAAGTTATCCAGAGTCTCCCTGTGTGAGAA 555
QY      400 AAATGTGACGCTTTCAGTGGGATTAATTTCTTTGGGAAGTGATAACGGTCGGAAACCC 459
Db      556 ACTTGTGACACATATTCTATGGTGTGTTCTCTGGGAGATGCTAAACAAGGAGGTCCCC 615
QY      460 TTTGATGAGATTGGTGG 476
Db      616 TTTAAAGGTTTGAAGG 632

```

Search completed: December 4, 2003, 07:06:36  
Job time : 89.7297 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2003, 04:21:38 ; Search time 730.77 seconds  
(without alignments)  
6784.777 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541

Perfect score: 204

Sequence: 1 caaagccgcgaaccttaacct.....agagcgtgtgacagcaccg 204

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*  
2: em\_esthum:\*\*  
3: em\_estin:\*\*  
4: em\_estmu:\*\*  
5: em\_estov:\*\*  
6: em\_estpl:\*\*  
7: em\_estro:\*\*  
8: em\_hic:\*\*  
9: gb\_est1:\*\*  
10: gb\_est2:\*\*  
11: gb\_hic:\*\*  
12: gb\_est3:\*\*  
13: gb\_est4:\*\*  
14: gb\_est5:\*\*  
15: em\_estfun:\*\*  
16: em\_estom:\*\*  
17: em\_gss\_hum:\*\*  
18: em\_gss\_inv:\*\*  
19: em\_gss\_pln:\*\*  
20: em\_gss\_vrt:\*\*  
21: em\_gss\_fun:\*\*  
22: em\_gss\_mam:\*\*  
23: em\_gss\_mus:\*\*  
24: em\_gss\_pro:\*\*  
25: em\_gss\_rod:\*\*  
26: em\_gss\_phg:\*\*  
27: em\_gss\_vrl:\*\*  
28: gb\_gss1:\*\*  
29: gb\_gss2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	204	100.0	757	13	BX393376
2	204	100.0	974	10	BE988567
3	204	100.0	1021	13	BUI52619
4	203	99.5	851	13	BX347763

5	200.8	98.4	782	12	BI253992
6	200.8	98.4	997	10	BGI15732
7	200.8	98.4	1385	13	BQ925425
8	194.4	95.3	930	13	BX347740
9	192.8	94.5	912	13	BX347764
10	184.8	90.6	934	13	BX347741
11	172	84.3	582	12	BI681846
12	165.6	81.2	378	10	BG382061
13	165.6	81.2	527	12	BI400511
14	165.6	81.2	599	12	BI401530
15	165.6	81.2	648	12	BI184403
16	165.6	81.2	815	12	BI181306
17	164.6	80.7	496	9	AM258218
18	164.6	80.7	578	13	BU609697
19	164.6	80.7	705	10	BE375985
20	164.6	80.7	777	12	BI181203
21	164	80.4	334	10	BF710376
22	160.2	78.5	644	12	BM934363
23	158	77.5	1107	10	BE902307
24	155.2	76.1	337	10	BF712308
25	153.6	75.3	738	13	BU709271
26	152.6	74.8	919	10	BF163305
27	145.8	71.5	325	9	AM658281
28	144	70.6	584	10	BE901639
29	141.2	69.2	818	10	BE746542
30	139.8	68.5	617	10	BG710962
31	139.8	68.5	772	13	BU229272
32	138.2	67.7	587	13	BU443120
33	138.2	67.7	941	13	BU427891
34	134.6	66.0	997	13	BX367448
35	130.4	63.9	677	13	BU426433
36	126.2	61.9	1002	14	BY709101
37	126.2	61.9	1576	11	AK009321
38	122.6	60.1	795	13	BU109063
39	120.4	59.0	450	9	AA674170
40	114.4	56.1	695	13	BU256845
41	113	55.4	654	13	BU417660
42	103	50.5	417	10	BE167965
43	102.4	50.2	614	12	BG913752
44	92.2	45.2	610	13	BQ387712
45	86.8	42.5	626	13	BX278260

## ALIGNMENTS

RESULT 1  
BX393376  
LOCUS  
DEFINITION  
BX393376 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DC002Y002 5-PRIME, mRNA sequence.  
ACCESSION  
BX393376  
VERSION  
BX393376.1 GI:30619996  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

757 bp mRNA linear EST 13-MAY-2003  
BX393376 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DC002Y002 5-PRIME, mRNA sequence.

REFERENCE  
1 (bases 1 to 757)  
AUTHORS  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10651.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAM001ZC05 AM029\_1&cluster=10651.f.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Faraday Avenue Genoscope sequence ID : CS0BAM001ZC05\_AM029\_1.  
Location/Qualifiers  
1. .757  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC002Y002"  
/tissue type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 162 a 253 c 197 g 145 t  
ORIGIN

Query Match 100.0%; Score 204; DB 13; Length 757;  
Best Local Similarity 100.0%; Pred. No. 1.4e-40;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTTCAGTCCACCAACACGACGACGAGGAGGAGCGGT 60  
Db 33 CAAAGCCGACCTTAACCTTCAGTCCACCAACACGACGACGAGGAGGAGCGGT 92  
QY 61 TCTGACGAGGCGCTTTCGCTCCCGCCCGCCGACCTCGCTCCCGCTCGCGAGGAGCGGT 120  
Db 93 TCTGACGAGGCGCTTTCGCTCCCGCCCGCCGACCTCGCTCCCGCTCGCGAGGAGCGGT 152  
QY 121 CGTGTGAGCCCTATGAGCTTTGCTGAGTTTACGCTCTGGAGCGTGGACCATGGC 180  
Db 153 CGTGTGAGCCCTATGAGCTTTGCTGAGTTTACGCTCTGGAGCGTGGACCATGGC 212  
QY 181 GAGCAGAGCGTGGTGACGACCG 204  
Db 213 GAGCAGAGCGTGGTGACGACCG 236

RESULT 2  
BE898567 974 bp mRNA linear EST 29-SEP-2000  
LOCUS 601681494F1 NIH\_MGC\_9 Homo sapiens cDNA clone IMAGE:3951886 5',  
DEFINITION mRNA sequence.  
ACCESSION BE898567  
VERSION BE898567.1 GI:10365176  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 974)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: L1CM820 row: k column: 23  
High quality sequence stop: 652.  
Location/Qualifiers  
1. .974  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3951886"  
/tissue type="adenocarcinoma cell line"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH\_MGC\_9"  
/note="Organ: ovary; Vector: pOTB7; Site\_1: XhoI; Site\_2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 209 a 311 c 278 g 176 t  
ORIGIN

Query Match 100.0%; Score 204; DB 10; Length 974;  
Best Local Similarity 100.0%; Pred. No. 1.4e-40;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTTCAGTCCACCAACACGACGACGAGGAGGAGCGGT 60  
Db 433 CAAAGCCGACCTTAACCTTCAGTCCACCAACACGACGACGAGGAGGAGCGGT 492  
QY 61 TCTGACGAGGCGCTTTCGCTCCCGCCCGCCGACCTCGCTCCCGCTCGCGAGGAGCGGT 120  
Db 493 TCTGACGAGGCGCTTTCGCTCCCGCCCGCCGACCTCGCTCCCGCTCGCGAGGAGCGGT 552  
QY 121 CGTGTGAGCCCTATGAGCTTTGCTGAGTTTACGCTCTGGAGCGTGGACCATGGC 180  
Db 553 CGTGTGAGCCCTATGAGCTTTGCTGAGTTTACGCTCTGGAGCGTGGACCATGGC 612  
QY 181 GAGCAGAGCGTGGTGACGACCG 204  
Db 613 GAGCAGAGCGTGGTGACGACCG 636

RESULT 3  
BU152619 1021 bp mRNA linear EST 03-SEP-2002  
LOCUS AGENCOURT\_8678094 NIH\_MGC\_40 Homo sapiens cDNA clone IMAGE:6381058  
DEFINITION 5', mRNA sequence.  
ACCESSION BU152619  
VERSION BU152619.1 GI:22666151  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1021)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: L1CM2570 row: k column: 11  
High quality sequence start: 48  
High quality sequence stop: 601.  
Location/Qualifiers  
1. .1021  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6381058"  
/tissue type="carcinoma, cell line"  
/lab host="DH10B (phage-resistant)"  
/clone lib="NIH\_MGC\_40"  
/note="Organ: prostate; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University

of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library." 6 others

BASE COUNT 205 a 386 c 259 g 165 t  
ORIGIN  
Query Match 100.0%; Score 204; DB 13; Length 1021;  
Best Local Similarity 100.0%; Pred. No. 1.4e-40;  
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAAGCCGAGCCTTAACCTGAGTCCACCAACGACGACGAGAGAGTCCAGC 60  
Db 244 CAAGCCGAGCCTTAACCTGAGTCCACCAACGACGACGAGAGTCCAGC 303  
QY 61 TCTGACGAGGAGCCTTCCGCTCCGCGCCGCCCACTCGCTCCGCTGGGAGACGGT 120  
Db 304 TCTGACGAGGAGCCTTCCGCTCCGCGCCGCCCACTCGCTCCGCTGGGAGACGGT 363  
QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGACCATGGC 180  
Db 364 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGACCATGGC 423  
QY 181 GAGCAGAGCGTGTGACGACCG 204  
Db 424 GAGCAGAGCGTGTGACGACCG 447

RESULT 4  
LOCUS BX347763 851 bp mRNA linear EST 05-MAY-2003  
DEFINITION BX347763 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
cDNA clone CS0DC018YG07 5-PRIME, mRNA sequence.  
ACCESSION BX347763  
VERSION BX347763.1 GI:30375210  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 851)  
AUTHORS Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10651.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAE008ZG12\_AE00768\_1&cluster=10651.f.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAE008ZG12\_AE00768\_1.  
Location/Qualifiers

FEATURES  
source  
1..851  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC018YG07"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 172 a 284 c 253 g 140 t  
ORIGIN  
Query Match 99.5%; Score 203; DB 13; Length 851;  
Best Local Similarity 99.5%; Pred. No. 2.4e-40;  
Matches 203; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAGCCGAGCCTTAACCTGAGTCCACCAACGACGACGAGAGAGTCCAGC 60  
Db 544 CAAGCCGAGCCTTAACCTGAGTCCACCAACGACGACGAGAGAGTCCAGC 603  
QY 61 TCTGACGAGGAGCCTTCCGCTCCGCGCCGCCCACTCGCTCCGCTGGGAGACGGT 120  
Db 604 TCTGACGAGGAGCCTTCCGCTCCGCGCCGCCCACTCGCTCCGCTGGGAGACGGT 663  
QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGACCATGGC 180  
Db 664 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGACCATGGC 723  
QY 181 GAGCAGAGCGTGTGACGACCG 204  
Db 724 GAGCAGAGCGTGTGACGACCG 747

RESULT 5  
LOCUS BI253992 782 bp mRNA linear EST 17-JUL-2001  
DEFINITION 602975130F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:5114622 5',  
mRNA sequence.  
ACCESSION BI253992  
VERSION BI253992.1 GI:14805965  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 782)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: csapbs@mail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM11278 row: k column: 07  
High quality sequence stop: 657.  
Location/Qualifiers  
1..782  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5114622"  
/tissue\_type="cervical carcinoma cell line"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_12"  
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.4 kb. Library prepared by Life  
Technologies."

BASE COUNT 162 a 288 c 214 g 118 t  
ORIGIN  
Query Match 98.4%; Score 200.8; DB 12; Length 782;  
Best Local Similarity 99.0%; Pred. No. 8.5e-40;  
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CAAGCCGAGCCTTAACCTGAGTCCACCAACGACGACGAGAGAGTCCAGC 60  
Db 306 CAAGCCGAGCCTTAACCTGAGTCCACCAACGACGACGAGAGAGTCCAGC 365  
QY 61 TCTGACGAGGAGCCTTCCGCTCCGCGCCGCCCACTCGCTCCGCTGGGAGACGGT 120  
Db 366 TCTGACGAGGAGCCTTCCGCTCCGCGCCGCCCACTCGCTCCGCTGGGAGACGGT 425

```

QY 121 CQTGTTGAGCCCTATGTGAGCTTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 180
    |||||||
Db 426 CQTGTTGAGCCCTATGTGAGCTTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 485
    |||||||
QY 181 GAGCAGAGCGTGGTGACAGCACCG 204
    |||||||
Db 486 GAGCAGAGCGTGGTGACAGCACCG 509
    |||||||

RESULT 6
BG115732
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 997)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10149 row: d column: 05
High quality sequence stop: 659.
Location/Qualifiers
1. .997
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4417108"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 88"
/notes="Organ: Small intestine; Vector: pCMV-SPORT6;
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
oligo-dt primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 251 a 310 c 296 g 140 t
ORIGIN

Query Match 98.4%; Score 200.8; DB 10; Length 997;
Best Local Similarity 99.0%; Pred. No. 8.9e-40;
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTCGAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 60
Db 194 CAATTCGCCGACTTAACCTCGAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 253
    |||||||
QY 61 TCTGACGAGGCGCTCTTCGCTCCCGCGCGCCGCTCGCTCCCGCTGGGAGGAGCGT 120
Db 254 TCTGACGAGGCGCTCTTCGCTCCCGCGCGCCGCTCGCTCCCGCTGGGAGGAGCGT 313
    |||||||
QY 121 CGTCTGAGCGCTATGTGAGCTTTGCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 180
Db 314 CGTCTGAGCGCTATGTGAGCTTTGCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 373
    |||||||
QY 181 GAGCAGAGCGTGGTGACAGCACCG 204
    |||||||
Db 374 GAGCAGAGCGTGGTGACAGCACCG 397
    |||||||

```

```

RESULT 7
BG925425
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1385)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
cDNA Library Preparation: ResGen, Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13734 row: c column: 07
High quality sequence start: 68
High quality sequence stop: 404.
Location/Qualifiers
1. .1385
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6310974"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 129"
/notes="Organ: olfactory epithelium; Vector:
pCMV-SPORT6.1.cdb; Site 1: EcoRV; Site 2: NotI; Cloned
unidirectionally. Primer: Oligo dT. Average insert size
2.2 kb. Constructed by ResGen, Invitrogen Corp. Note: this
is a NIH_MGC Library."
BASE COUNT 257 a 536 c 337 g 253 t 2 others
ORIGIN

Query Match 98.4%; Score 200.8; DB 13; Length 1385;
Best Local Similarity 99.0%; Pred. No. 9.5e-40;
Matches 202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CAAAGCCGACCTTAACCTCGAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 60
Db 247 CAAAGCCGACCTTAACCTCGAGTCCACCAACGACGACGAGCAGCAGCTCCAGC 306
    |||||||
QY 61 TCTGACGAGGCGCTCTTCGCTCCCGCGCGCCGCTCGCTCCCGCTGGGAGGAGCGT 120
Db 307 TCTGACGAGGCGCTCTTCGCTCCCGCGCGCCGCTCGCTCCCGCTGGGAGGAGCGT 366
    |||||||
QY 121 CGTCTGAGCGCTATGTGAGCTTTGCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 180
Db 367 CGTCTGAGCGCTATGTGAGCTTTGCTGAGTTTACCGCTCTGAGCGTGGACCATGGC 426
    |||||||
QY 181 GAGCAGAGCGTGGTGACAGCACCG 204
    |||||||
Db 427 GAGCAGAGCGTGGTGACAGCACCG 450
    |||||||

RESULT 8
BX347740
LOCUS
DEFINITION
ACCESSION
VERSION

```

```

BX347740 930 bp mRNA linear EST 05-MAY-2003
BX347740 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DC002Y002 5--PRIME, mRNA sequence.
BX347740
BX347740.1 GI:30369230

```

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL 1 (bases 1 to 910)  
COMMENT Full-length cDNA libraries and normalization  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Unpublished

CONTACT: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10651.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAE007ZE10\_AE00654\_1&cluster=10651.f.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAE007ZE10\_AE00654\_1.  
Location/Qualifiers

FEATURES  
source  
1..930  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC002Y002"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 168 a 316 c 224 g 219 t 3 others  
ORIGIN

Query Match 95.3%; Score 194.4; DB 13; Length 930;  
Best Local Similarity 97.1%; Pred. No. 3.4e-38;  
Matches 198; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CAAGCCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 60  
Db 429 CAAGCCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 488

QY 61 TCTGACGAGGCGCTTTCGCTCCGCGCCGACCTCGCTCCGCTCGGAGCGTGGACCATGGC 120  
Db 489 TCTGACGAGGCGCTTTCGCTCCGCGCCGACCTCGCTCCGCTCGGAGCGTGGACCATGGC 548

QY 121 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 180  
Db 549 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 608

QY 181 GACGAGCGTGTGACAGCACCG 204  
Db 609 GACGAGCGTGTGACAGTACCG 632

RESULT 9  
BX347764  
LOCUS  
DEFINITION  
BX347764 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
cDNA clone CS0DC018YG07 5-PRIME, mRNA sequence.  
ACCESSION  
VERSION  
BX347764.1 GI:30375211  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL 1 (bases 1 to 912)  
COMMENT Full-length cDNA libraries and normalization  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Unpublished

COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10651.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAE008ZG12\_AE00768\_2&cluster=10651.f.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAE008ZG12\_AE00768\_2.  
Location/Qualifiers

FEATURES  
source  
1..912  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC018YG07"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 181 a 298 c 269 g 159 t 5 others  
ORIGIN

Query Match 94.5%; Score 192.8; DB 13; Length 912;  
Best Local Similarity 96.6%; Pred. No. 8.4e-38;  
Matches 197; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CAAGCCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 60  
Db 542 CAAGCCCGACCTTAACCTTGCAGTCCACCAACGACGACGACGACGACGACG 601

QY 61 TCTGACGAGGCGCTTTCGCTCCGCGCCGACCTCGCTCCGCTCGGAGCGTGGACCATGGC 120  
Db 602 TCTGACGAGGCGCTTTCGCTCCGCGCCGACCTCGCTCCGCTCGGAGCGTGGACCATGGC 661

QY 121 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 180  
Db 662 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCCTCTGGAGCGTGGACCATGGC 721

QY 181 GACGAGCGTGTGACAGCACCG 204  
Db 722 GACGAGCGTGTGACAGCACCG 745

RESULT 10  
BX347741  
LOCUS  
DEFINITION  
BX347741 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
cDNA clone CS0DC002Y002 5-PRIME, mRNA sequence.  
ACCESSION  
VERSION  
BX347741.1 GI:30369231  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL 1 (bases 1 to 934)  
COMMENT Full-length cDNA libraries and normalization  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Unpublished

CONTACT: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10651.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAE007ZE10\_AE00654\_2&cluster=10651.f.

Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAP007ZB10\_AE00654\_2.

FEATURES  
source  
1..934  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC002Y002"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 164 a 308 c 205 g 253 t 4 others  
ORIGIN

Query Match 90.6%; Score 184.8; DB 13; Length 934;  
Best Local Similarity 94.1%; Pred. No. 8.1e-36;  
Matches 192; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 1 CAAGCCGACCTTAACCTCGAGTCCACCAACACGACACGAGAGGAGGAGCTCCAGC 60  
DB 430 CAAGCCGACCTTAACCTCGAGTCCACCAACACGACACGAGAGGAGGAGCTCCAGC 489  
QY 61 TCTGACGGAGGCGCTTTCGGCTCCCGGCCCGCCCACTCGCTCCCGCTCGGAGGAGCGGT 120  
DB 490 TCTGACGGAGGCGCTTTCGGCTCCCGGCCCGCCCACTCGCTCCCGCTCGGAGGAGCGGA 549  
QY 121 CGTGTGACCCCTATGTGACCTTTCGTGAGTTTACCGCTCTGAGCGGTGGACCATGGC 180  
DB 550 CGTGATGAGCCCTATGTGACCTTTCGTGAGTTTACCGCTCTGAGCGGTGGTCCATGGC 609  
QY 181 GAGCAGGCGTGTGACAGCAGCG 204  
DB 610 GTGACAGCGTGTGACAGTTCGG 533

RESULT 11  
BI681846 582 bp mRNA linear EST 17-SEP-2001  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Sequence evaluation of four pooled-tissue normalized bovine cDNA  
libraries and construction of a gene index for cattle  
Genome Res. 11 (4), 626-630 (2001)  
21180013  
11282978  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCAGTCCAGCAGC  
Plate: 135 row: 0 column: 15  
Seq primer: ATTTAGGTGACACTATAG.  
Location/Qualifiers  
1..582  
/organism="Bos taurus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9913"  
/tissue\_type="pooled"  
/lab\_host="DH10B"  
/clone\_lib="MARC 1B0V"  
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;  
Library made from pooled tissue from lymph node, ovary,  
fat, hypothalamus, and pituitary."  
BASE COUNT 93 a 212 c 179 g 98 t  
ORIGIN

Query Match 84.3%; Score 172; DB 12; Length 582;  
Best Local Similarity 90.2%; Pred. No. 1.1e-32;  
Matches 184; Conservative 0; Mismatches 20; Indels 0; Gaps 0;  
QY 1 CAAGCCGACCTTAACCTCGAGTCCACCAACACGACACGAGAGGAGGAGCTCCAGC 60  
DB 80 CAAGCCGACCTTAACCTCGAGTCCACCAACACGACACGAGAGGAGGAGCTCCAGC 139  
QY 61 TCTGACGGAGGCGCTTTCGGCTCCCGGCCCGCCCACTCGCTCCCGCTCGGAGGAGCGGT 120  
DB 140 TCGACGGGGCGCTTTCGGCTCCCGGCCCGCCCACTCGCTCCCGCTCGGAGGAGTGC 199  
QY 121 CGTGTGACCCCTATGTGACCTTTCGTGAGTTTACCGCTCTGAGCGGTGGACCATGGC 180  
DB 200 CGGTGGAGCCCTACGTGACCTTTCGGAGTTTACCGCTCTGAGCGGTGGACCATGGC 259  
QY 181 GAGCAGGCGTGTGACAGCAGCG 204  
DB 260 GAGCAGGCGTGTGACGGCGCG 283

RESULT 12  
BG382061 378 bp mRNA linear EST 12-MAR-2001  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

297803 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.  
BG382061  
BG382061.1 GI:13306533  
EST.  
Sus scrofa (pig)  
Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
Fahrenkrug, S.C., Smith, T.P.L., Preking, B.A., Cho, J., White, J.,  
Vallet, J., Wise, T., Rohrer, G.A., Pertea, G., Sultana, R., Quackenbush,  
J. and Keele, J.W.  
Porcine gene discovery by normalized cDNA-library sequencing and  
EST cluster assembly  
Mamm. Genome 13 (8), 475-478 (2002)  
22213789  
12226715  
Contact: Smith TPL  
USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -minscore 18  
and -mismatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCTCAGTCCAGCAGC  
Plate: 2 row: 0 column: 16  
Seq primer: ATTTAGGTGACACTATAG.  
Location/Qualifiers

FEATURES



source

```

1. .378
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/lab_host="DH10B"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC iPig"
/notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT      52 a 154 c 123 g 49 t
ORIGIN

Query Match      81.2%; Score 165.6; DB 10; Length 378;
Best Local Similarity 88.2%; Pred. No. 3.9e-31;
Matches 180; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 CAAGCCGACCTTAACCTGAGTCCACACGACACGACGAGCAGAGCTCCAGC 60
Db 71 CAGAGCCGACCTGACCTGAGTCCACACACACACACACGAGCAGCTCCAGC 130
QY 61 TCTGACGAGGCTCTCCGCTCCGCGCCGCGCCGCTGCTCCGCGCGGAGACCGT 120
Db 131 TCCGAGGGGGCTCTTTCGCTCCGCGCCGCGCCGCTGCTCCGCGCGGAGATGC 190
QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTTCTGGAGCGTGACCATGC 180
Db 191 CGCGTGAGGCTTACGTGAGTCTCGCGAGTTCTACCGCTTCTGGAGCGTGACCATGC 250
QY 181 GAGCAGAGCGTGTGACGACCG 204
Db 251 GAGCAGAGTGTGGCGCGCGCG 274

RESULT 13
BI400511/c
LOCUS      BI400511      527 bp      mRNA      linear      EST 14-AUG-2001
DEFINITION MI-P-Ayl-nge-a-03-0-UI.s1 MI-P-Ayl Sus scrofa cDNA clone
ACCESSION  BI400511
VERSION     BI400511.1 GI:15179572
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1 (bases 1 to 527)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      889548
COMMENT     Contact: Tuggle CK
            Molecular Genetics Laboratory, Department of Animal Science
            Iowa State University
            201 Kilgus Hall, Ames, IA 50011-3150, USA
            Tel: 5152944252
            Fax: 5152942401
            Email: cktuggle@iastate.edu
            Oligo-dr track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab,
            University of Iowa Clone distribution: clones will be available
            through Research Genetics (www.resgen.com) The following repetitive
            elements were found in this cDNA sequence: 62-112,
            >GC-rich#Low_complexity
            Seq primer: M13 Forward
            POLVA-No.

FEATURES             Location/Qualifiers
source               1. .527
                    /organism="Sus scrofa"
                    /mol_type="mRNA"

```

```

/strain="crossbreed"
/db_xref="taxon:9823"
/clone="MI-P-Ayl-nge-a-03-0-UI"
/lab_host="DH10B (Life Technologies)"
/clone_lib="MI-P-Ayl"
/notes="Vector: pTT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: EcoRI; The MI-P-Ayl
library is normalized library derived from the MI-P-Ayl
library, ultimately derived from placenta tissue. For a
detailed description of the library from which this clone
was derived, please visit our web site at
http://pigest.genome.iastate.edu/. The procedure used to
create this library has been previously described (Bonaldo
, Lennon and Soares, Genome Research 6: 791-806, 1996)
TAG_SEQ=None found"
BASE COUNT      67 a 166 c 216 g 78 t
ORIGIN

Query Match      81.2%; Score 165.6; DB 12; Length 527;
Best Local Similarity 88.2%; Pred. No. 4.2e-31;
Matches 180; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 1 CAAGCCGACCTTAACCTGAGTCCACACGACACGACGAGCAGAGCTCCAGC 60
Db 318 CAGAGCCGACCTGACCTGAGTCCACACACACACACGAGCAGAGCTCCAGC 259
QY 61 TCTGACGAGGCTCTTTCGCTCCGCGCCGCGCCGCTGCTCCGCGCGGAGACCGT 120
Db 258 TCCGAGGGGGCTCTTTCGCTCCGCGCCGCGCCGCTGCTCCGCGCGGAGATGC 199
QY 121 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTTCTGGAGCGTGACCATGC 180
Db 198 CGCGTGAGGCTTACGTGAGTCTCGCGAGTTCTACCGCTTCTGGAGCGTGACCATGC 139
QY 181 GAGCAGAGCGTGTGACGACCG 204
Db 138 GAGCAGAGTGTGGCGCGCGCG 115

RESULT 14
BI401530/c
LOCUS      BI401530      599 bp      mRNA      linear      EST 14-AUG-2001
DEFINITION MI-P-CP0-nvn-g-07-0-UI.s1 MI-P-CP0 Sus scrofa cDNA clone
ACCESSION  BI401530
VERSION     BI401530.1 GI:15180591
KEYWORDS    EST.
SOURCE      Sus scrofa (pig)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1 (bases 1 to 599)
AUTHORS     Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE       Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL     Genome Res. 6 (9), 791-806 (1996)
MEDLINE     97044477
PUBMED      889548
COMMENT     Contact: Tuggle CK
            Molecular Genetics Laboratory, Department of Animal Science
            Iowa State University
            201 Kilgus Hall, Ames, IA 50011-3150, USA
            Tel: 5152944252
            Fax: 5152942401
            Email: cktuggle@iastate.edu
            Oligo-dr track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab,
            University of Iowa Clone distribution: clones will be available
            through Research Genetics (www.resgen.com) The following repetitive
            elements were found in this cDNA sequence: 60-110,
            >GC-rich#Low_complexity
            Seq primer: M13 Forward

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 08:22:54 ; Search time 23.2027 Seconds  
(without alignments)  
4537.638 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541

Perfect score: 361

Sequence: 1 caaagccgacctaaccct.....agagcgtggtgacagcaccg 204

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh  
-Q=/cgn2 1/USPTO.spool/US09830144/runat\_03122003\_122344\_21277/app\_query.fasta\_1.1230  
-DB=SPTRMBL\_23 -CPMP=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 @CGN 1 1 100 @runat\_03122003\_122344\_21277 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	356	98.6	500	11	Q8R0D1	Q8r0d1 mus musculus

ID	Q8R0D1	PRELIMINARY;	PRT;	500 AA.
2	349	96.7	502	11
3	275	76.2	52	11
4	260.5	72.2	498	13
5	85	22.6	340	4
6	82	21.8	189	2
7	80.5	21.4	625	11
8	79.5	21.1	501	11
9	78.5	20.9	531	10
10	78	21.6	464	12
11	78	20.7	6193	2
12	77.5	20.6	356	4
13	77	20.5	404	4
14	76	20.2	268	11
15	75.5	20.1	349	2
16	74.5	20.6	329	4
17	74.5	19.8	336	4
18	74	19.7	111	10
19	74	19.7	252	10
20	74	19.7	273	10
21	73.5	20.4	311	10
22	73.5	19.5	369	10
23	73	19.4	335	12
24	73	20.2	402	10
25	73	19.4	563	10
26	73	19.4	1039	5
27	72.5	19.3	70	6
28	72.5	20.1	233	2
29	72.5	20.1	532	12
30	72.5	19.3	947	10
31	72	19.9	138	4
32	72	19.1	310	12
33	72	19.1	383	12
34	72	19.1	383	12
35	72	19.1	383	12
36	72	19.1	383	12
37	72	19.1	383	12
38	72	19.1	383	12
39	72	19.1	383	12
40	72	19.1	383	12
41	72	19.1	383	12
42	72	19.1	437	12
43	72	19.1	500	16
44	72	19.9	574	3
45	71.5	19.0	467	10

## ALIGNMENTS

RESULT 1  
Q8R0D1  
ID Q8R0D1 PRELIMINARY; PRT; 500 AA.  
AC Q8R0D1;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Similar to mitogen-activated protein kinase kinase 7  
DE Interacting protein 1 (Fragment).  
GN 2310012M03RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC027054; AAH27054.1; -  
DR MGD; MGI:1913763; 2310012M03RIK.  
DR InterPro; IPR001932; PP2C-like.  
DR Pfam; PF00481; PP2C; 1.  
DR SMART; SM00332; PP2CC; 1.  
KW Kinase.

Q8cf89 mus musculus  
Q8cv62 mus musculus  
Q73614 xenopus lae  
Q96h27 homo sapien  
Q91730 streptomyce  
Q99053 rattus norv  
Q8ci00 rattus norv  
Q94e81 oryza sativ  
Q91tm2 tupaiia herp  
Q8ksq0 streptomyce  
Q8iw96 homo sapien  
Q96b18 homo sapien  
Q8cdk3 mus musculus  
Q47925 frankia sp.  
Q96s04 homo sapien  
Q92623 homo sapien  
Q39682 daucus caro  
Q8la16 arabidopsis  
Q9m9r2 arabidopsis  
Q9awn1 oryza sativ  
Q94ll0 oryza sativ  
Q91889 spodoptera  
Q8m655 oryza sativ  
Q8ryj3 oryza sativ  
Q9n974 leishmania  
Q9xs8r canis fami  
Q8gges streptomyce  
Q40622 human papil  
Q91wt9 oryza sativ  
Q9nv39 homo sapien  
Q69130 human herpe  
Q04397 epstein-bar  
Q8azk8 human herpe  
Q8azk6 human herpe  
Q8azk5 human herpe  
Q8azk4 human herpe  
Q8azk3 human herpe  
Q8azk2 human herpe  
Q8azk1 human herpe  
Q69146 human herpe  
Q9rxk6 streptomyce  
Q36027 schizosacch  
Q9fwp6 oryza sativ





Thu Dec 4 17:00:26 2003

us-09-830-144-3\_copy\_1338\_1541.rspt

RT "Functional Analysis of the Profilaggrin N-Terminal Peptide:  
RT Identification of Domains that Regulate Nuclear and Cytoplasmic  
RT Distribution";  
RT J. Invest. Dermatol. 119:661-669(2002).  
DR EMBL; AY102923; AAM54369.1; -.  
FT NON TER 501 501  
SQ SEQUENCE 501 AA; 55418 MW; 7317F95F907FD235 CRC64;

QY 28 ---TGAGCTGCAGGG 17  
Db 57 ArgTrpThrSerGly 61

RESULT 7  
Q99053 PRELIMINARY; PRT; 625 AA.

AC Q99053; 1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Profilaggrin (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90274870; PubMed=1693512;  
RA Haydock P.V., Dale B.A.;  
RT "Filaggrin, an intermediate filament-associated protein: Structural  
RT and functional implications from the sequence of a cDNA from rat.";  
RL DNA Cell Biol. 9:251-261(1990).  
DR EMBL; M21759; AAA41161.1; -.  
FT NON TER 1 1  
FT CHAIN 1 406 FILAGGRIN.  
SQ SEQUENCE 625 AA; 65957 MW; 8424342E121CB105 CRC64;

Alignment Scores:  
Pred. No.: 2.48 Length: 625  
Score: 80.50 Matches: 28  
Percent Similarity: 47.44% Conservative: 9  
Best Local Similarity: 35.90% Mismatches: 27  
Query Match: 21.41% Indels: 14  
DB: 11 Gaps: 3

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x Q99053 (1-625)

QY 196 TCACACGCTGCTCGCATGTCACGCTCCAGAGCGGTAAACT-----CAGCAA 143  
Db 116 ThrProGlySerThrArgThrGlySerArgGlyGluSerProAlaGlyGlnGln 135  
QY 142 AGTCCACATAGGCTCAACAC-----GACCGTCTCGCCAGCGGAGCG 98  
Db 136 SerProAspArgAlaArgHisIleGluSerArgGlyArgThrArgGluAlaSerAla 155  
QY 97 AGT-----GGCGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 56  
Db 156 SerGlnSerSerAspSerGluGlyHisSerGlyAlaHisAlaGlyIleGlyGlnGln 175  
QY 55 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4  
Db 176 Thr-SerThrIleHisArgArgAlaGlySerSerSerGlySerGlnArgAla 192

RESULT 8  
Q8CIU0 PRELIMINARY; PRT; 501 AA.

AC Q8CIU0; 2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Profilaggrin (Fragment).  
GN FLG.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Skin;  
RX MEDLINE=2218272; PubMed=12230510;  
RA Pearson D.J., Dale B.A., Presland R.B.;

RT "Functional Analysis of the Profilaggrin N-Terminal Peptide:  
RT Identification of Domains that Regulate Nuclear and Cytoplasmic  
RT Distribution";  
RT J. Invest. Dermatol. 119:661-669(2002).  
DR EMBL; AY102923; AAM54369.1; -.  
FT NON TER 501 501  
SQ SEQUENCE 501 AA; 55418 MW; 7317F95F907FD235 CRC64;

Alignment Scores:  
Pred. No.: 3.15 Length: 501  
Score: 79.50 Matches: 24  
Percent Similarity: 47.46% Conservative: 4  
Best Local Similarity: 40.68% Mismatches: 14  
Query Match: 21.14% Indels: 17  
DB: 11 Gaps: 3

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x Q8CIU0 (1-501)

QY 196 TCACACGCTGCTCGCATGTCACGCTCCAGAGCGGTAAACT-----CAGCAA 143  
Db 295 ThrProGlySerThrArgThrGlySerArgGlyGluSerProAlaGlyGlnGln 314  
QY 142 AGTCCACATAGGCTCAACAC-----GACCGTCTCGCCAGCGGAGCG 98  
Db 315 SerProAspArgAlaArgHisIleGluSerArgGlyArgThrArgGluAlaGlyAla 334  
QY 97 AGT-----GGCGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCG 71  
Db 335 SerGlnSerSerAspSerGluAlaIleSerGlyAlaHisAlaGlyIleGlyGlnGly 353

RESULT 9  
Q94E81 PRELIMINARY; PRT; 531 AA.

AC Q94E81; 2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Putative 3-dehydroquinase dehydratase.  
GN B1111C09.4.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC  
RT clone:B1111C09.4";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003204; BAB61159.1; -.  
DR Gramene; Q94E81; -.  
DR InterPro; IPR001381; DHQuinase\_I.  
DR InterPro; IPR000205; NAD\_binding.  
DR InterPro; IPR006152; Shikimate.  
DR InterPro; IPR006151; Shikimate\_DH.  
DR Pfam; PF01487; DHQuinase\_I; 1.  
DR Pfam; PF01488; Shikimate\_DH; 1.  
DR ProDom; PD005337; DHQuinase\_I; 1.  
DR TIGRFAMs; TIGR01093; arod; 1.  
DR TIGRFAMs; TIGR00507; aroe; 1.  
SQ SEQUENCE 531 AA; 56905 MW; 17E3491DC42CE95A CRC64;

Alignment Scores:  
Pred. No.: 4.07 Length: 531  
Score: 78.50 Matches: 23  
Percent Similarity: 46.34% Conservative: 15  
Best Local Similarity: 28.05% Mismatches: 23  
Query Match: 20.88% Indels: 21  
DB: 10 Gaps: 4

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x Q94E81 (1-531)

```

Alignment Scores:
Pred. No.:      4.58
Score:          78.00
Percent Similarity: 52.27%
Best Local Similarity: 47.73%
Query Match:    21.61%
DB:             12
Length:         464
Matches:        21
Conservative:   2
Mismatches:     19
Indels:         2
Gaps:           1

US-09-830-144-3_COPY_1338_1541 (1-204) x Q91TM2 (1-464)

QY      8 CGACCTTAACTCCGTTCAGTCCACCACAACAGCAGCACGCGAGCAGCTCCAGCTCTCAAG 67
Db      421 ArgProArgProArgProArgProArgProAlaAlaAAlaProThrArgSer 440
QY      68 GAGCGCTCTTCCGCTCCCGGCCGCC-----ACTGGCTCCCGCTGGCGAGCAGCGTC 121
Db      441 ArgAlaglyTyrSerProGlyProProProArgThrArgSerSerThrArgArgValVal 460
QY      122 GTGTTTGAGCCCT 133
Db      461 ValSerSerPro 464

RESULT 11
Q8KSQO

```







***This Page Blank (uspro)***

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: December 4, 2003, 08:37:54 ; Search time 8.95946 Seconds  
(without alignments)  
4379.374 Million cell updates/sec

Title: US-09-830-144-3\_COPY\_1338\_1541  
Perfect score: 361  
Sequence: 1 caaagccgcgaaccttaacct.....agagcgtggtgacagcaccg 204

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh  
-Q/cgn2 1/USPTO spool/US09830144/runat\_03122003\_122345\_21308/app\_query.fasta\_1.1230  
-DB=pir 76 -QWMT=fasan -SUFFIX=rpr -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=-1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US09830144 @CGN 1 1 35 @runat\_03122003\_122345\_21308 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR 76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	80.5	21.4	625	A34615	profilaggrin - rat
C 2	74	19.7	111	T14306	glycine-rich prote
C 3	74	19.7	273	A86280	F14L17-29 protein
C 4	73	20.2	518	P2WL47	L2 protein - human
C 5	72	19.1	383	S32975	gene BCRF2 protein
C 6	72	19.9	574	T43556	Wiskott-Aldrich sy
C 7	72	19.9	574	T38819	wiskott-aldrich sy
C 8	72	19.9	1199	A40670	nuclear envelope p
C 9	71	19.7	132	S15618	E4 protein - human
C 10	71	18.9	152	T34649	hypothetical prote
C 11	71	18.9	219	B84326	hypothetical prote
C 12	71	19.7	240	G83401	hypothetical prote
C 13	71	19.7	519	S36471	L2 protein - human
C 14	71	19.7	521	S36571	L2 protein - human

15	70.5	19.5	518	2	S36542	L2 protein - human
C 16	70	18.6	142	2	T00616	probable nucleic a
C 17	70	18.6	256	2	A84437	probable PHD-type
C 18	70	19.4	520	2	S36489	L2 protein - human
C 19	70	18.6	606	2	G75302	orotidine 5'-phosp
C 20	69.5	18.5	177	2	C87614	hypothetical prote
C 21	69.5	18.5	1146	2	A38587	collagen, cornea-s
C 22	69.5	19.3	1198	2	T28678	polyketide synthas
C 23	68.5	18.2	537	2	B46535	interleukin 2 rece
C 24	68.5	18.2	824	2	B38423	protein-glutamine
C 25	68.5	18.2	993	2	F84774	probable ATP-depen
C 26	68	18.1	190	2	C72532	hypothetical prote
C 27	68	18.1	511	2	S24345	Balbani ring 1 pr
C 28	68	18.8	667	2	S63587	gene pacc protein
C 29	68	18.1	1428	2	T13926	probable protein p
C 30	68	18.1	1666	2	T09072	probable translati
C 31	67.5	18.7	125	1	A46315	E4 protein - human
C 32	67.5	18.0	386	2	A41950	retrovirus-related
C 33	67.5	18.7	704	2	S21911	BRcore-NS-23 prote
C 34	67	17.8	682	2	T28899	hypothetical prote
C 35	67	18.6	989	2	C83035	hypothetical prote
C 36	67	18.6	1083	2	T05689	hypothetical prote
C 37	66.5	17.7	257	2	T09646	probable zinc fing
C 38	66.5	18.4	550	2	T36746	probable serine/th
C 39	66	18.3	384	2	JC5255	TEA domain-contain
C 40	66	18.3	427	2	JC5254	L2 protein - human
C 41	66	18.3	518	1	P2WL8	L2 protein - human
C 42	66	18.3	518	1	P2WL5	protein-glutamine
C 43	66	17.6	817	1	TGHUM1	hypothetical orf5
C 44	65.5	17.4	102	4	JC5161	ORF2 protein - Orf
C 45	65.5	17.4	221	2	C34768	

#### ALIGNMENTS

##### RESULT 1

A34615

profilaggrin - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-Jun-1990 #sequence\_revision 09-Oct-1992 #text\_change 04-Feb-2000

C:Accession: A34615

R:Haydock, P.V.; Dale, B.A.

DNA Cell Biol. 9, 251-261, 1990

A:Title: Filaggrin, an intermediate filament-associated protein: structural and functions

A:Reference number: A34615; MUID:90274870; PMID:1693512

A:Accession: A34615

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-625 <HAY>

A:Cross-references: GB:M21759; NID:G204143; PIDN:AAA41161.1; PID:G204144

A>Note: the authors translated the codon GAA for residue 568 as Gln

C:Superfamily: unassigned calmodulin-related proteins; calmodulin repeat homology

C:Keywords: EF hand; epidermis

Alignment Scores:

Pred. No.:	0.826	Length:	625
Score:	80.50	Matches:	28
Percent Similarity:	47.44%	Conservative:	9
Best Local Similarity:	35.90%	Mismatches:	27
Query Match:	21.41%	Indels:	14
DB:	2	Gaps:	3

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x A34615 (1-625)

QY 196 TCACACAGCTCTGCTCGCATGGTCCACGTCAGAGCGGTAAACT-----CAGCAA 143

Db 116 ThrProGlySerThrArgThrGlySerArgThrGlySerArgThrProAlaGlyGlnGln 135

QY 142 AGTCACATAGAGGCTCAACAC-----GACCGTCCTCCAGCGGAGCGG 98

Db 136 SerProAspArgAlaArgHisIleGluSerArgGlyArgThrArgGluAlaSerAla 155

QY 97 AGT-----GGGCGGCGCGGAGCGGAAGAGCGCTCGTCGTCAGAGCTCG 56



RESULT 6  
T43556  
Wiskott-Aldrich syndrome protein homolog - fission yeast (Schizosaccharomyces pombe)  
C.Species: Schizosaccharomyces pombe  
C.Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jan-2000  
C.C.Accession: T43556  
R.Zankel, T.C.; Ow, D.W.  
Submitted to the EMBL Data Library, December 1997  
A.Description: A Wiskott-Aldrich Syndrome protein homolog in Schizosaccharomyces pombe,  
A.Reference number: Z22575  
A.Accession: T43556  
A.Status: preliminary; translated from GB/EMBL/DBDJ  
A.Molecule type: DNA  
A.Residues: 1-574 <ZAN>  
A.Cross-references: EMBL:AF038575; PIDN:AAB92587.1  
A.Experimental source: strain JS21  
C.Genetics:  
A.Gene: wsp1  
A.Map position: I  
A.Introns: 72/3; 519/3; 564/1

**A;Title:** An integral membrane protein of the pore membrane domain of the nuclear envelope

A:Reference number: A40670; MUID:93328754; PMID:8335683  
A:Accession: A40670  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1199 <HAL>  
A:Cross-references: GB:Z21513; NID:9396746; PIDN:CAA79725.1; PID:9396747  
F1803-807,845-849,956-960,1010-1014,1047-1051,1076-1080/Region: pentapeptide motif (X-F-  
Alignment Scores:  
Pred. No.: 7.16 Length: 1199  
Score: 72.00 Matches: 19  
Percent Similarity: 47.37% Conservative: 8  
Best Local Similarity: 33.33% Mismatches: 18  
Query Match: 19.94% Indels: 12  
DB: 2 Gaps: 1

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x A40670 (1-1199)  
QY 17 CCTGAGTCCACACACGACGACGAGTCCAGCTCTGACGGAGGCTCT 76  
Db 757 ProLeuSerProPheSerLeuLysGlnThrThrAlaProAlaThrThrAlaAlaThr 776  
QY 77 TCGCTCCGGCCGGCCCACTCGCTCCGCTGGCGAGGACGCTCGTGTGTGAGCCCTATG 136  
Db 777 SerAlaPro-----LeuLeuThrGlyLeu 784  
QY 137 TGGACTTCTGAGTTTACCGCTCTGAGCGTGGACCATGGCGAGCAGA 187  
Db 785 GlyThrAlaThrSerThrValAlaThrGlyThrThrAlaSerAlaSerLys 801

RESULT 9  
S15618  
E4 protein - human papillomavirus type 2a  
C:Species: human papillomavirus type 2a  
A:Note: host Homo sapiens (man)  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 17-Feb-1994  
C:Accession: S15618  
R:Hirsch-Behnam, A.; Delius, H.; de Villiers, E.M.  
Virus Res. 18, 81-98, 1990  
A:Title: A comparative sequence analysis of two human papillomavirus (HPV) types 2a and  
A:Reference number: S15614; MUID:91188699; PMID:1964523  
A:Accession: S15618  
A:Molecule type: DNA  
A:Residues: 1-132 <HIR>  
A:Cross-references: EMBL:X55964  
C:Superfamily: papillomavirus type 2 E4 protein  
C:Keywords: early protein

Alignment Scores:  
Pred. No.: 9.87 Length: 132  
Score: 71.00 Matches: 22  
Percent Similarity: 45.45% Conservative: 3  
Best Local Similarity: 40.00% Mismatches: 20  
Query Match: 19.67% Indels: 10  
DB: 1 Gaps: 3

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x S15618 (1-132)  
QY 11 CCTTAACCTTCGAGTCCACCA-----ACACGCACACGACGACGAGCTCCAGCTCTG 64  
Db 66 ProLysArgCysAlaProProArgGlnArgValArgProSerAlaSerValSer 85  
QY 65 ACGGAGGCTCTTCGGCTCCGGCCCGCCACATCGCTCCGCTGGCGGAGGACGGTCTG 124  
Db 86 SerSerAspSerSerLysProGlyProThrLeuArgGluArgSerGluArg----- 102  
QY 125 TTGAGCCCTATGGACTTTGCTGAGTTTACCGCTCTGGAGCG 169  
Db 103 -----GlyLysTrp-----SerValThrThrSerGlyAla 112

RESULT 10  
T34649  
hypothetical protein SC1A11.02c SC1A11.02c - Streptomyces coelicolor (fragment)

C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
A:Accession: T34649  
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, January 1999  
A:Reference number: Z21551  
A:Accession: T34649  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-152 <SEE>  
A:Cross-references: EMBL:AL035205; PIDN:CAA22742.1; GSPDB:GN00070; SCOEDB:SC1A11.02c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC1A11.02c

Alignment Scores:  
Pred. No.: 9.83 Length: 152  
Score: 71.00 Matches: 24  
Percent Similarity: 46.97% Conservative: 7  
Best Local Similarity: 36.36% Mismatches: 29  
Query Match: 18.88% Indels: 6  
DB: 3 Gaps: 3

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x T34649 (1-152)  
QY 196 TCACACGCTCTGCTCGCATGCTCCAGCTCCAGAGCGGTAAACTCAGCAAGTCCA 137  
Db 38 SerProGlyThrGlyArgAlaargGlyArgAlaargGlyArg-----Pro 53  
QY 136 CATAGGCGCTCAACACAGACCGCTCTCGCCAGCGGGAGCGAGTGGCGGCGCGG---AGC 80  
Db 54 TrpArgArgSerAspSerThrCysGlySerAlaGlyCysSerArgTrpThrGlyArgThr 73  
QY 79 GGAAGAGCGCTCGCTCAGAGCTGG---AGCTGCTGCTCTGCGTGGTGGTGGTGGACT 23  
Db 74 GlyArgAlaGlyProArgSerTrpArgThrCysThrThrAlaGlyCysCysArgArgThr 93  
QY 22 GCAGGGTTAAGGTCGGGC 5  
Db 94 IleSerArgArgSerGly 99

RESULT 11  
B84326  
hypothetical protein Vng1743c [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: B84326  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.;  
Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabin,  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L.  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: AB4160; MUID:20504483; PMID:11016950  
A:Accession: B84326  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-219 <STO>  
A:Cross-references: GB:AE004437; NID:910581201; PIDN:AAG19974.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: VNG1743C

Alignment Scores:  
Pred. No.: 9.73 Length: 219  
Score: 71.00 Matches: 26  
Percent Similarity: 36.25% Conservative: 3  
Best Local Similarity: 32.50% Mismatches: 23  
Query Match: 18.88% Indels: 28  
DB: 2 Gaps: 4

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x B84326 (1-219)  
QY 196 TCACACGCTCTGCTCGCAT-----GGTCCAGCTCCAGAGCGGTAAACTCAGCAA 143

Db 11 SerProArgSerThrArgThrArgCysGlyArgArgSerArgArgThrProThrAlaArg 30  
QY 142 AGTCCACATAGGGCTCAACACAGCGTCTCGCCAGCGGAGCGAGTGGCGGCGCGG 83  
Db 31 ArgArgHisArgGlyValGlnProArgPro-----ThrAlaArgArgProThr 46  
QY 82 AGCGGAAGAGCGCTCCGCTCAGAGCTGAGCTGCTCTTGGCGTGGCGGT----- 32  
Db 47 ThrGlyArg-----ArgProSerTrpSerArgValSerValMetAlaValArgThrLeu 64  
QY 31 -----TGGTGGACT 23  
Db 65 ProArgArgArgArgTrpAlaAlaProThrArgAlaArgArgSerArgArgTrpThr 84  
RESULT 12  
GB3401  
hypothetical protein PA1960 [imported] - Pseudomonas aeruginosa (strain PAO1)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: GB3401  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho  
A:Reference number: AB2950; MUID:2043737; PMID:10984043  
A:Accession: GB3401  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-240 <STO>  
A:Cross-references: GB:AE004622; GB:AE004091; NID:g9947948; PIDN:AA05348.1; GSPDB:GN001  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA1960

Alignment Scores:  
Pred. No.: 9,7 Length: 240  
Score: 71.00 Matches: 21  
Percent Similarity: 50.00% Conservative: 5  
Best Local Similarity: 40.38% Mismatches: 24  
Query Match: 19.67% Indels: 2  
DB: 2 Gaps: 2

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x GB3401 (1-240)  
QY 26 CCACCAACAGCGACACGAGACAGCAGTCCAGCTCTGACGGAGGCGCTCTCGGTCCC 85  
Db 86 ProProThrValArgArgGlyAlaGlnAlaPro---ValThrArgAlaGlnAlaArgPro 104  
QY 86 GGCCCGCCCATCTGCTCCGCTCGCGTGGCGAGGAGCGTGTG---TTGAGCCCTATGTGGACT 142  
Db 105 AlaProThrProProSerAlaAlaGlnArgProLeuAlaLysLeuValLeuLeu 124  
QY 143 TTGCTGAGTTTTACGCGCTCTGGAGCGTGGACCATG 178  
Db 125 LeuCysValLeuProAlaAlaGlyAlaTrpArgMet 136  
RESULT 13  
S36471  
L2 protein - human papillomavirus type 14D  
C:Species: human papillomavirus type 14D  
C:Date: 09-Dec-1993 #sequence\_revision 26-Jul-1996 #text\_change 26-Aug-1999  
C:Accession: S36471  
R:DeLius, H.; Hofmann, B.  
submitted to the EMBL Data Library, August 1993  
A:Description: Primer-directed sequencing of human papillomavirus types.  
A:Reference number: S36469  
A:Accession: S36471  
A:Molecule type: DNA  
A:Residues: 1-519 <DEL>  
A:Cross-references: EMBL:X74467; NID:g396918; PIDN:CAA52504.1; PID:g396922  
A:Experimental source: strain 14D

Thu Dec 4 17:00:26 2003

A;Cross-references: EMBL:X74466; NID:G396910; PIDN:CAA52500.1; PID:G396916  
 C;Superfamily: papillomavirus L2 protein  
 C;Keywords: late protein

Alignment Scores:  
 Pred. No.: 10.8 Length: 518  
 Score: 70.50 Matches: 20  
 Percent Similarity: 56.36% Conservative: 11  
 Best Local Similarity: 36.36% Mismatches: 21  
 Query Match: 19.53% Indels: 3  
 DB: 2 Gaps: 2

US-09-830-144-3\_COPY\_1338\_1541 (1-204) x S36542 (1-518)

QY	23	AGTCCACCAACACGACGACGACGACGCTCTGACGAGCGCTCTTCCGCT	82
Db	150	SerProValValThrThrSerArgGlySerSer---AlaIleuGluValAlaProAsp	168
QY	83	CCCGGCCCGCCACTCGCTCCCGCTCGCGAGGACGCTGTGTGAGCCCTATGTGGACT	142
Db	169	ProIleProThrArgValArgValAlaArgThrGlnTyrHisAsnProAlaPheGln	188
QY	143	TTGCTG-----AGTTTACCGCTCTGGAGCGTGGACCATGGCG	181
Db	189	IleIleThrGluSerThrProAlaGlnGlyGluThrSerIleuAla	203

Search completed: December 4, 2003, 09:08:41  
 Job time : 11.9595 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: December 4, 2003, 09:42:22 ; Search time 2899.3 seconds  
(without alignments)  
1911.298 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303  
Perfect score: 1252  
Sequence: 1 VELQLSRVNHPIVLYGA.....MTHLMRYPGADEPLQVPCQ 228

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09830144/runat\_03122003\_122509\_23332/app\_query.fasta\_1.654  
-DB=EST -QPM=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOPFCI=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09830144 @CGN 1 1 3596 @runat\_03122003\_122509\_23332 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estm:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gesl:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	100.0	2994	11	AK053911 Mus muscu
2	1246	99.5	1221	13	BUS11353 AGENCOURT
3	1236	98.7	2976	11	AK046565 Mus muscu
4	1167	93.2	739	12	B1696710 603347214
5	1099	87.8	1201	13	BX364931 BX364931
6	1069	85.4	604	12	BM426610 BM426610
7	1041	83.1	839	13	BU418819 BU418819
8	1024	81.8	1157	14	CD508913 CD508913
9	1016	81.2	1062	12	BM554120 BM554120
10	998	79.7	750	13	BU104161 BU104161
11	984	78.6	646	12	BJ073883 BJ073883
12	982	78.4	688	12	BJ062988 BJ062988
13	982	78.4	696	12	BJ074867 BJ074867
14	966	77.2	600	14	CB112561 CB112561
15	953.5	76.2	910	10	BF780358 BF780358
16	932.5	74.5	1087	9	AL525728 AL525728
17	924	73.8	590	14	CB583299 CB583299
18	849.5	67.9	1118	9	AL550589 AL550589
19	847	67.7	542	10	AW960377 AW960377
20	824	65.8	668	13	BUS29310 BUS29310
21	823	65.7	641	13	BQ386875 BQ386875
22	802.5	64.1	742	13	BQ058580 BQ058580
23	753	60.1	1006	13	BQ219348 BQ219348
24	748	59.7	686	10	BG548917 BG548917
25	741	59.2	771	12	BI093821 BI093821
26	737	58.9	682	9	AV398933 AV398933
27	723	57.7	924	14	CA984799 CA984799
28	715	57.1	586	12	BJ035129 BJ035129
29	707	56.5	681	9	AV398935 AV398935
30	705	56.3	1016	9	AL520975 AL520975
31	695	55.5	707	13	BW233644 BW233644
32	663	53.0	894	14	CA792459 CA792459
33	631.5	50.4	1054	13	BQ067693 BQ067693
34	631	50.4	607	9	AI541706 AI541706
35	628	50.2	594	12	BJ014293 BJ014293
36	613	49.0	588	13	BU381175 BU381175
37	604	48.2	609	9	AL860068 AL860068
38	587	46.9	747	12	BI334774 BI334774
39	559	44.6	516	12	BJ006787 BJ006787
40	557	44.5	858	12	BG916821 BG916821
41	553	44.2	711	9	AV718168 AV718168
42	552	44.1	630	13	BW238127 BW238127
43	552	44.1	710	13	BX473160 BX473160
44	549.5	43.9	760	12	BG773488 BG773488
45	549	43.8	822	12	BI757369 BI757369

# ALIGNMENTS

RESULT 1  
AK053911

LOCUS

DEFINITION

AK053911.1

GI:26343862

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

AK053911  
Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30320K22 product:mitogen activated protein kinase kinase kinase 7, full insert sequence.

2994 bp mRNA linear HTC 05-DEC-2002  
Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30320K22 product:mitogen activated protein kinase kinase kinase 7, full insert sequence.

AK053911.1  
GI:26343862  
HTC; CAP trapper.  
Mus musculus (house mouse)  
Mus musculus  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

## JOURNAL

99279253  
MEDLINE  
10349636  
PUBMED

## REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

20499374  
MEDLINE  
11042159  
PUBMED

## REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

20530913  
MEDLINE  
11076861  
PUBMED

## REFERENCE

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Harai, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Staehli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hall, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyo-oka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.,  
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

21085660  
MEDLINE  
11217851  
PUBMED

## REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

6 (bases 1 to 2994)

## JOURNAL

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Muramatsu, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome

EXPLORATION RESEARCH GROUP, RIKEN GENOMIC SCIENCES CENTER (GSC),  
RIKEN YOKOHAMA INSTITUTE: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
KANAGAWA 230-0045, JAPAN (E-mail: genome-res@gs.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

Location/Qualifiers  
1..2994  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:EL30320K22"  
/db\_xref="taxon:10090"  
/clone="EL30320K22"  
/tissue\_type="eyeball"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="0 day neonate"  
143..1882  
/note="unnamed protein product; mitogen activated protein  
kinase kinase kinase 7 (MGD|MG1:1346877, GB|D76446,  
evidence: BLASTN, 99%, match=2403)  
putative"  
/codon\_start=1  
/protein\_id="BAC355588.1"  
/db\_xref="GI:26343863"  
/translation="MSTASASSSSSSSSEMEAPSOVLNFEIDYKEIEVEEIVGR  
GAFVCKAKRAKVAIKQIESERKAFIVELRSLSRVNHVNIKLYGACINPVCL  
NLLVAGGTVLKICDGTACDIQTHMNNKSGAAMWAVEFESNYSSEKCDVFSWGLI  
LMEVITRRKPFDEIGGPAFRIMWANGTRPPLIKNLPKIESLMTRCWSKDSQRP  
MEIVKIMTHLMRYPPGAPQYPCQYDEGOSNATSTGSMFDTASTNSKSDTM  
MQQVPAINTIKRLSKLLKNAQKQSSRLSAGSGSVESLPTTSGTKSEKAP  
SEIEARIVATAGNGQRRRSIQDLTVITGPGQVSSRSPPSVMTTSGPTSEKAP  
SHWTPDDSTDTGNSDINSIPMAYLTLDHQIQLAPCPNSKESMAVFEQCKMAQEVNK  
VOTEIALLKQKQELVAELDQDEKQDQNTSRIVQEHKKLLDENKSLSTYFQCKKQLE  
VIRSQQKQKQTS"

BASE COUNT 854 a 657 c 749 g 734 t

ORIGIN

Alignment Scores:  
Pred. No.: 4,07e-124 Length: 2994  
Score: 1252.00 Matches: 228  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 11 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x AK053911 (1-2994)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyGlyVala 20

368 GTGAGCTCCGCGAGTGTGTGACGATGCAATTCCTAACATTTCAAGTTGTATGAGACC 427

QY 21 CysLeuAsnProValCysLeuValMetGluTyAlaGluGlyGlySerLeuTyAsnVal 40

428 TGCCTGAATCCAGTATGCTCTGTGATGGAATATGACAGGGGGGCTCATTTGATAATGTG 487

QY 41 LeuHisGlyAlaGluProLeuProTyTyTyThrAlaAlaHisAlaMetSerTrpCysLeu 60

488 CTGCATGCTGCTGAACCATTCCTCTACTACATGCTCTCATGCGATGAGTGTGTTA 547

QY 61 GlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProIysAlaLeuIleHisArg 80

548 CAGTGTTCCCAAGGAGTGCTTACTGACAGCATGACGCCAAGCGCTGATTCACAGG 607

QY 81 AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100

```

Db      608  GACCTCAAGCCTCCAAACTTGCTGCTGGTTGCAGGAGGACAGTTCTAAATATCGCGAT 567
Qy      101  PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db      668  TTGTTGATCAGCTTGATCATCCAAACACACATGACCAATAATAAAGGGAGTGTCTGCTGG 727
Qy      121  MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db      728  ATGGCGCCTGAAGTATTGTAAGGTAGCAATTACAGTGAAGGTGATGTCTTCAGCTGG 787
Qy      141  GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160
Db      788  GGTATTATPCTCTGGGAAGTGTATACACCGCGGAAACCTTCGATGAGATCGGTGGCCCA 847
Qy      161  AlaPheArgIleMetTrpAlaValHisGlnGlyThrArgProProLeuIleLysAsnLeu 180
Db      848  GCTTTCAGAAATCATGTGGGCTGTTTCAATATGGCACTCGACCACTGATCAAAAATTTA 907
Qy      181  ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db      908  CCTAAGCCCAATTGAGAGCTTGATGACACGCTGTGGTCTAAGGACCCATCTCAGCGCCCT 967
Qy      201  SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db      968  TCAATGGAGGAAATTTGAAATAATGACTCACTTGATGCGGTACTTCCAGGAGCGGAT 1027
Qy      221  GluProLeuGlnTyrProCysGln 228
Db      1028  GAGCCGTTACAGTATCTCTTGTCAG 1051

RESULT 2
BUS11353      1221 bp      mRNA      linear      EST 12-SEP-2002
LOCUS      AGENCOURT 10128981 NIH_MGC_134 Mus musculus cDNA clone
DEFINITION      IMAGE:6506145 5', mRNA sequence.
ACCESSION      BUS11353
VERSION      BUS11353.1 GI:22817586
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-k@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLAM14068 row: c column: 10
High quality sequence stop: 733.
FEATURES      Location/Qualifiers
source      1..1221
            /organism="Mus musculus"
            /mol_type="mRNA"
            /db_xref="taxon:10090"
            /clone="IMAGE:6506145"
            /tissue_type="undifferentiated limb"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC 134"
            /note="Vector: pCMV-SPORT6.1.cdb; Site 1: EcoRV; Site 2:
            NotI. Cloned unidirectionally. Primer: Oligo dt. Average
            insert size 1.7 kb. Constructed by ResGen, Invitrogen
            Corp. Note: this is a NIH MGC Library."
BASE COUNT      337 a 287 c 319 g 277 t 1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      4.51e-124      Length:      1221
Score:      1246.00      Matches:      227
Percent Similarity:      99.56%      Conservative:      0
Best Local Similarity:      99.56%      Mismatches:      1
Query Match:      99.52%      Indels:      0
DB:      13      Gaps:      0

US-09-830-144-2_COPY_76_303 (1-228) x BUS11353 (1-1221)

Qy      1  ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db      148  GTGAGACTCCGCGAGTTGTCGCGTGTGAACCATCTTAACATTTGTCAAGTTGTACGAGCC 207
Qy      21  CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db      208  TGCCTGAATCCAGTATGTTCTTGATGAATATGAGAGGGGGCTCAATTGTATATATGTG 267
Qy      41  LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
Db      268  CTGCATGCTGTGAACCATTCCTTACTACACTGCTCTCATGCCATGAGCTGGTGTTA 327
Qy      61  GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db      328  CAGTGTTCCTCCCAAGGAGTGGCTTACCTGCACAGCATGCAGCCCAAGCGCTGATTCACAGG 387
Qy      81  AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db      388  GACCTCAAGCCTCCAAACTTGTCTGCTGTCAGAGGGAGCAGTTCTAAATAATCTGGAT 447
Qy      101  PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db      448  TTTGGTACAGCTTGTGACATCCAAACACACATGACCAATAATAAAGGGAGTGTCTGTGG 507
Qy      121  MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db      508  ATGCGCCTCAAGTGTGTTGAAGGTAGCAATTCAGTGAAGAGTGTGATGCTTCAGCTGG 567
Qy      141  GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db      568  GGTATTATCTCTGGGAAGTGTATACACGCGGAAACCTTCGATGAGATCGGTGGCCCA 627
Qy      161  AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db      628  GCTTTCAGATCATGTGGGCTGTTCAATATGGCACTGCACCACTGATCAAAAATTTA 687
Qy      181  ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db      688  CCTAAGCCCAATTGAGAGCTTGATGACACGCTGTGGTCTAAGGACCCATCTCAGCGCCCC 747
Qy      201  SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db      748  TTCATGGAGAAATTTGAAATAATATGACTCACTTGTGCGGTACTTCCAGGAGCGGAT 807
Qy      221  GluProLeuGlnTyrProCysGln 228
Db      808  GAGCCATTACAGTATCTCTTGTCAG 831

RESULT 3
AK046565      2976 bp      mRNA      linear      HTC 05-DEC-2002
LOCUS      Mus musculus 4 days neonate male adipose cDNA, RIKEN full-length
DEFINITION      enriched library, clone:B430101B05 product:mitogen activated
              protein kinase kinase kinase 7, full insert sequence.
ACCESSION      AK046565
VERSION      AK046565.1 GI:26091589
KEYWORDS      HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1

```

## AUTHORS

Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

## MEDLINE

99279253  
PUBMED 10349636

## REFERENCE

## AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

## JOURNAL

20493374  
PUBMED 11042159

## REFERENCE

## AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

## TITLE

## JOURNAL

20530913  
MEDLINE 11076861

## PUBMED

## REFERENCE

## AUTHORS

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamauaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavante, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hall, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohetsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

## TITLE

## JOURNAL

21085660  
MEDLINE 11217851

## PUBMED

## REFERENCE

## AUTHORS

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

## TITLE

## JOURNAL

6 (bases 1 to 2976)

## REFERENCE

## AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, I., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

## TITLE

## JOURNAL

767

## COMMENT

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/.

## FEATURES

## Location/Qualifiers

## Source

1..2976  
/organism="Mus musculus"  
/mol\_type="rRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:B430101B05"  
/db\_xref="taxon:10090"  
/clone="B430101B05"  
/sex="male"  
/tissue\_type="adipose"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="4 days neonate"  
124..1862  
/note="mitogen activated protein kinase kinase 7 (MGD) [MGI:1346877, GB|D76446, evidence: BLASTN, 99%, match=2403]  
putative"  
2963..2968  
/note="putative"  
2976  
/note="putative"  
850 a 851 c 743 g 732 t  
BASE COUNT  
ORIGIN

## misc\_feature

## polyA\_signal

## polyA\_site

## Alignment Scores:

Pred. No.: 2,17e-122 Length: 2976  
Score: 1236.00 Matches: 226  
Percent Similarity: 99.12% Conservative: 0  
Best Local Similarity: 99.12% Mismatches: 2  
Query Match: 98.72% Indels: 0  
DB: 11 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x AK046565 (1-2976)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyVala 20

Db 348 GTGGAGCTCGGCAGTTGTCCACGTGTGAACCATCTTAACATTGTCAAGTTGTATGGAGCC 407

Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40

Db 408 TGCGTGAATCCAGTATGTCGTGATGGGATATGAGAGGGGGCTCATTTGATTAATGTG 467

Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60

Db 468 CTGCATGGTGTGACCATTTGCCTTACTACATCTCTCTCATGCCATGCGTGTGTTTA 527

Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80

Db 528 CAGTGTTCCTCAAGGAGTGGCTTACCTGCACAGCATGCGCCCAAGGCTGATTCACAGG 587

Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100

Db 588 GACCTCAAGCCTCCAACTTCTCTGTTGTCAGAGGGGACAGTTCTAAAAATCTCGCAT 647

Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrrp 120

Db 648 TTTCGTACAGCTTGTGACATCCAAACACACATGACCAATAATAAAGGAGTGTCTCTGG 707

Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrrp 140

Db 708 ATGGCGCATGAAGTATTGGAAGGTAGCAATTACAGTGAAGAGTGTGTCTTTCAGCTGG 767

```

QY 141 GlylleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
DB 768 GGTATTATCTCTGGGAAGTGATAACACCGCGGAACCTTTCGATGAGATCGTGGCCCA 827
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
DB 828 GCTTTTCAAGATCATGTGGCTCTTATATGCGACTCGACACCACTGATCAAAAATTATA 887
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 888 CCTAGACCAATTCAGAGCTTGTATGACACGCTGTTGCTTAAGGACCCATCTCAGGCGCT 947
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyPheProGlyAlaAsp 220
DB 948 TCAATAGGAGGAATGTGAAATAATGACTCACTTGTGCGGTACTTCCAGGACGGAT 1007
QY 221 GluProLeuGlnTyProCysGln 228
DB 1008 GAGCGTTACAGTATCTTGTTCAG 1031

RESULT 4
LOCUS BI696710 739 bp mRNA linear EST 18-SEP-2001
DEFINITION 603347214F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374971 5',
mRNA sequence.
ACCESSION BI696710
VERSION BI696710.1 GI:15659339
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11954 row: k column: 04
High quality sequence stop: 717.
FEATURES
    source
        1..739
            /organism="Mus musculus"
            /mol_type="mRNA"
            /strain="FVB/N-3"
            /db_xref="taxon:10090"
            /clone="IMAGE:5374971"
            /tissue_type="tumor, biopsy sample"
            /dev_stage="5 months"
            /lab_host="DH10B"
            /clone_lib="NCI_CGAP_Mam2"
            /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
            Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
            Library constructed by Life Technologies. Investigator
            providing samples: Gilbert Smith, NIH"
BASE COUNT 195 a 169 c 188 g 187 t
ORIGIN

Alignment Scores:
Pred. No.: 7.35e-116 Length: 739
Score: 1167.00 Matches: 220
Percent Similarity: 98.65% Conservative: 0
Best Local Similarity: 98.65% Mismatches: 1
Query Match: 93.21% Indels: 3
DB: 12 Gaps: 0

```

```

US-09-830-144-2_COPY_76_303 (1-228) x BI696710 (1-739)
QY 8 ArgValAsnHisProAsnIleValLysLeuTyGlyAlaCysLeuAsnProValCysLeu 27
DB 3 CGTGTGAACCATCTTAACATTCAGTTGTACGGAGCCTGCCTGAATCCAGTATGCTT 62
QY 28 ValMetGluTyTrpAlaGluGlySerLeuTyAsnValLeuHisGlyAlaGluProLeu 47
DB 63 GTGATGGAATATGCAGAGGGGGCTCATTTATATATGCTGTCATGTGCTGACCATG 122
QY 48 ProTyTrpThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly-Valal 67
DB 123 CCTTACTACCTGCTGCATGCCATGAGCTGGTGTTTACAGTGTTCACAGGTAGTGC 182
QY 67 aTyTrpLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProAsnLe 87
DB 183 TTACCTGCACAGCATGCAGCCCAAGCGCTGATTCACAGGACCTCAAGCCTCCAACTT 242
QY 87 uLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAspI 107
DB 243 GCTGCTGGTTGCAGAGGAGGACAGTTCTAAAAATCTCGGATTTTGGTACAGCTTGTG 302
QY 107 eGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheG 127
DB 303 CCAACACACATGACCAATAATAAGGGAGTGTCTTGGATGGCGCTGAAGTGTGTA 362
QY 127 uGlySerAsnTySerGluLysCysAspValPheSerTrpGlyIleIleLeuTyGluVa 147
DB 363 AGGTAGCAATTACAGTGAAGAGTGTGATCTTCAGCTGGGTATTATCTCTGGGAAGT 422
QY 147 lIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAl 167
DB 423 GATAACACGCGCGAAACCTTCGATGAGATCGGTGGCCAGCTTTCAGATCATGTGGGC 482
QY 167 aValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLe 187
DB 483 TGTTCAATAATGGCACTCGACCACTGATCAAAAATTTACCTAAGCCATTGAGAGCTT 542
QY 187 uMetThrArgCys-TrpSerLysAspProSerGlnArgProSerMetGluGluIleVal 207
DB 543 GATGACACGCTGTGTGCTTAAGACCCATCTCAGGC-CCTTCAATGGAGGAATTTGTA 601
QY 207 YstIleMetThrHisLeuMetArgTy-PheProGlyAlaAspGluProLeuGlnTyProC 227
DB 602 AAATAATGACTCACTTGTATGCGTACTTCCAGAGGCGGATGAGCCATTACAGTATCCTT 661
QY 227 YsGln 228
DB 662 GTCAG 666

RESULT 5
LOCUS BX364931 1201 bp mRNA linear EST 05-MAY-2003
DEFINITION BX364931 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CS0DI032YB01 5-PRIME, mRNA sequence.
ACCESSION BX364931
VERSION BX364931.1 GI:30376963
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3613.r For

```

more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSIAI008ZH01QP1&cluster=3613.r>. Contact :  
 Feng liang Email : fliang@lifetech.com URL :  
<http://fulllength.invitrogen.com/> InvitroGen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CSIAI008ZH01QP1.

FEATURES  
 Location/Qualifiers  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS01032YB01"  
 /tissue type="PLACENTA COT 25-NORMALIZED"  
 /clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="First strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT6 vector. Library was normalized."  
 BASE COUNT 290 a 253 c 320 g 299 t 39 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,58e-108 Length: 1201  
 Score: 1099.00 Matches: 213  
 Percent Similarity: 95.95% Conservative: 0  
 Best Local Similarity: 95.95% Mismatches: 9  
 Query Match: 87.78% Indels: 3  
 DB: 13 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x BX364931 (1-1201)  
 QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20  
 Db 446 GTAGAGCTTCGGCAGTATCCCGTGTGAACCACTCTAATATGTAAGCTTTATGAGCC 505  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluClyGlySerLeuTyrAsnVal 40  
 Db 506 TGCTTGAATCCAGTGCTGTCTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 565  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
 Db 566 CTCGATGGTGTGAACCATTTGCCATATATCTGCTGCCACGCAATGATGTTGTTA 625  
 QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80  
 Db 626 CAGTGTTCACAGAGAGTGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 685  
 QY 81 AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIleCysAsp 100  
 Db 686 GACCTGAACACCAACCAACTTACTGCTGTCGCGGGGGACAGTCTTAAATAATTTGTAT 745  
 QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTyr 120  
 Db 746 TTTGGTACAGCTGTGACATTCACACACATGACCAATACCAAGGGAGTGCTGCTGG 805  
 QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluIysCysAspValPheSerTyr 140  
 Db 806 ATGGCACTGAAGTTTGAAGGTAGTAATATACAGTGAATAATGACGCTCTTCAGCTGG 865  
 QY 141 GlyIleLeuLeuTyrGluValIleThrArgArgIysProPheAspGluIleGlyClyPro 160  
 Db 866 GGTATTATCTTCGGGAAGTGATACCGCTCGGAACCCCTTGTAGATGTTGGTGCCCA 925  
 QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleGlyAsnLeu 180  
 Db 926 GCTTTCGAAACATGTGGCTGCTCATATGTAATGTCGACCACTGATATAAAAAATTA 985  
 QY 181 ProIysProIleGluSerLeuMetThrArgCysTyrSerIysAspProSerGluArgPro 200  
 Db 986 CCTAAGCCCATGAGAGCGCTGATGACTGCTGTGTTGGTCTAAGATCCCTTCCAGCG-CCY 1044  
 QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220

Db 1045 TCAATGGAGGAATTGTGAAAAATAATGTCMC-TTGATGGCGGACTTTCM-GGGAGAGAT 1102  
 QY 221 GluPro 222  
 Db 1103 GAGCCW 1108  
 RESULT 6  
 LOCUS BM426610 604 bp mRNA linear EST 30-JAN-2002  
 DEFINITION pgf2n.pk004.d15 Normalized Chicken Abdominal Fat Library (pgf2n)  
 Gallus gallus cDNA clone pgf2n.pk004.d15 5' similar to  
 emb|CAB87605.1 (AL121964) dU154G14.1.4 [mitogen-activated protein  
 kinase kinase kinase 7 (TGF-beta activated kinase 1d (TAK1))] (Homo  
 sapiens), mRNA sequence.  
 ACCESSION BM426610  
 VERSION BM426610.1 GI:18431087  
 KEYWORDS EST.  
 SOURCE Gallus gallus (chicken)  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 REFERENCE 1 (bases 1 to 604)  
 AUTHORS Cogburn,L.A., Morgan,R. and Burnside,J.  
 TITLE ESTs from Normalized Chicken fat cDNA library-USDA/IFAFS Animal  
 JOURNAL Genome Project  
 COMMENT Unpublished  
 CONTACT: Larry A. Cogburn  
 University of Delaware  
 Townsend Hall, Newark, DE 19717, USA  
 Tel: 302-831-1335  
 Fax: 302-831-2822  
 Email: cogburn@udel.edu, www.chickest.udel.edu.  
 FEATURES  
 source  
 1..604  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Commercial broiler, Ottawa Research Centre,  
 leghorn"  
 /db\_xref="taxon:9031"  
 /clone="pgf2n.pk004.d15"  
 /sex="Male and Female"  
 /tissue type="Abdominal Fat"  
 /dev stage="Embryonic (dl8,dl9); post-hatch (dl,w3,w7,w9  
 w16,tyr)"  
 /lab\_host="E. coli BMDH10B"  
 /clone lib="Normalized Chicken Abdominal Fat Library  
 (pgf2n)"  
 /notes="Vector: pCMVSPORT6; Library made from equivalent  
 pools of total RNA isolated from each developmental age  
 (across strains); Single pass sequencing from 5'-end"  
 BASE COUNT 167 a 125 c 140 g 167 t 5 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.17e-105 Length: 604  
 Score: 1069.00 Matches: 194  
 Percent Similarity: 98.99% Conservative: 2  
 Best Local Similarity: 97.98% Mismatches: 0  
 Query Match: 85.38% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-830-144-2\_COPY\_76\_303 (1-228) x BM426610 (1-604)  
 QY 30 GluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGluProLeuProTyr 49  
 Db 2 GAGTATGCTGAAGGAGGTTCTGTACATGTGTGTGATGTGTGACCTCTGCTCAT 61  
 QY 50 TyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGlyValAlaTyrLeu 69  
 Db 62 TATACTGCTGCACACGCAATGAGTTGGTTTACAGTGTTCACAGGAGTGATATCTT 121  
 QY 70 HisSerMetGlnProIysAlaLeuIleHisArgAspLeuIysProProAsnLeuLeuLeu 89

```

Db 122 CACAGTATGAACCAAGAGCCCTTAATTCACAGAGACCTGAACACCAATTTGCTCTTG 181
Qy 90 ValAlaGlyGlyThrValLeuLysLeuCysAspPheGlyThrAlaCysAspIleGlnThr 109
Db 182 GTAGCTGGGGGACAGTCTTAAAGATCTGTGAATTTGGTACAGCCTGTGATATCAACA 241
Qy 110 HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer 129
Db 242 CACATGACCAACAATAAGGAAGTCTGCTTGGATGGACCTGAAGTTTGAAGGTAGC 301
Qy 130 AsnTyrSerGluLysCysAspValPheSerTrpGlyIleLeuLeuLeuLeuValIleThr 149
Db 302 AATTACAGTGAATAATGTGACGCTTTTCAGTTGGGTATTTATCTTTGGAGGTAATCACC 361
Qy 150 ArgArgLysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaValHis 169
Db 362 CGTAGGAACCTTTTGATGAGATGGTGTCCAGCTTTCCGCATAATGTGGGCAGTTCCAC 421
Qy 170 AsnGlyThrArgProProLeuLysAsnLeuProLysProLysProLysProLysProLysMetThr 189
Db 422 AATGGTACTCGACCACTGATCAAAACCTTACTTAACCAATTGAGAGTTTATGACC 481
Qy 190 ArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db 482 CGCTGTGTGGTCCAAAGATCCCTCAACACGACCTTCCATGGAGGAAATTTGTTAAATAATG 541
Qy 210 ThrHisLeuMetArgTyrPheProGlyValAlaAspGluProLeuGlnTyrProCys 227
Db 542 ACACACTTGATCGGTACTTTCNNNAGCTGATGAACCTCTGCAGTATCCTTGC 595

```

```

RESULT 7
BU418819 839 bp mRNA linear EST 29-NOV-2002
LOCUS 603962410F1 CSEQRN09 Gallus gallus cDNA clone CHEST939d3 5', mRNA
DEFINITION
ACCESSION
VERSION
SOURCE
ORGANISM

```

```

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomedical Sciences
University of Manchester Institute of Science and Technology (UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

```

```

FEATURES
source
1..839
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="CHEST939d3"
/sex="Male and female"
/tissue type="Chondrocytes isolated from growth plate cartilage"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="CSQRN09"
/note="Vector: pBluescript II KS(+); Site_1: EcoRI;

```

Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT 237 a 182 c 205 g 215 t

ORIGIN

Alignment Scores:

Pred. No.: 3.89e-102 Length: 839  
 Score: 1041.00 Matches: 196  
 Percent Similarity: 98.51% Conservative: 3  
 Best Local Similarity: 97.03% Mismatches: 0  
 Query Match: 83.15% Indels: 3  
 DB: 13 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x BU418819 (1-839)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyVala 20

Db 232 GTAGAGCTTCGACAACTGTCCAGAGTAACCAATCCCAATATTGTCAAGTATATATGGAGCC 291

Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40

Db 292 TGTCTTAACACCAAGTGTCTTGTATGGAGTAGCTGCTGAGGAGGTTCTCTGTACAATGTG 351

Qy 41 LeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60

Db 352 CTGCATGGTCTGAACCTCTGCCTCATTTACTGCTGCACACCAATGAGTTGGTGTGTTA 411

Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80

Db 412 CAGTGTTCACCAAGAGTGGCATATCTTCAGATATGAACCAAGGCCCTTAATTCACAGA 471

Qy 81 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100

Db 472 GACCTGAAACCAACCAATTTGCTTGTGAGTGGGGGACAGTCTCTAAAGATCTGTGAT 531

Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120

Db 532 TTTGGTACGCTGTGTATTTCAACACACATGACCACCAATTAAGGAGTGTGCTTGG 591

Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140

Db 592 ATGGCACCTGAAGTTTTCAGAGGTAGCAATTCAGTGAATAATGTGACGTTTTCACGTGG 651

Qy 141 GlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIle-GlyGlyPr 160

Db 652 GGTATTATTCTTTGGGAGGTAATCACCGTAGAAACCTTTTGTATGAGATGGTGGTCTTC 711

Qy 160 oAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLe 180

Db 712 AGCTTTCCGCATAATGTGGGCAGTTCACAATGGTATCTGCACCACCACTGATCAAAACTT 771

Qy 180 uProLysProIleGluSerLeu-MetThrArgCysTrpSerLysAsp-ProSerGlnArg 199

Db 772 ACCTAAACCAATTGAGAGTTTAAATGACCCGCTGTGTGTCGAAGGATCCCTCCACAAAGG 831

RESULT 8

CD508913

LOCUS

DEFINITION

CD508913

ACCESSION

VERSION

1157 bp mRNA linear EST 12-JUN-2003  
 CDA93-E03.x1d-t SHGC-CDA Gasterosteus aculeatus cDNA clone  
 CDA93-E03 5', mRNA sequence.  
 CD508913  
 CD508913.1 GI:31439482

KEYWORDS	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Gasterosteus aculeatus (three spined stickleback)	Gasterosteus aculeatus	1. (bases 1 to 1157)	Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M., Schmutz, J., and Myers, R.M.	Expressed sequence tags from <i>Gasterosteus aculeatus</i>	Unpublished	Contact: Kingsley, DM HHMI and Department of Developmental Biology Stanford University School of Medicine Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA Tel: 650 725 5954 Fax: 650 725 7739 Email: kingsley@cmsgm.stanford.edu Plate: 93 High quality sequence stop: 796.
FEATURES	source	1. .1157	Location/Qualifiers	/organism="Gasterosteus aculeatus"	/mol_type="mRNA"	/strain="Salinas river, CA"
				/db_xref="taxon:69293"	/clone="CDA93-E03"	/sex="mixed male and female"
				/tissue type="heads and internal organs combined"	/dev stage="adult"	/clone_lib="SHGC-CDA"
				/note="Vector: lambda ZAP Express/pBK-CMV; Site 1: EcoRI (5' adaptor); Site 2: XhoI (3' linker primer); The mixed organ cDNA library was generated using the ZAP-cDNA method by Stratagene. First strand cDNA synthesis was primed with a 50 bp linker primer containing an oligo dT sequence preceded by a synthetic XhoI site. 5 prime adaptors were used containing an EcoRI cohesive end. The finished cDNAs were inserted in to the ZAP express vector		
						undirectionally in the sense orientation with respect to the lacZ promoter of pBK-CMV. An amplified library was prepared from approximately 3 million primary clones in the lambda ZAP Express vector. In vivo excision was then used to generate individual pBK-CMV phagemid clones for EST sequencing."
BASE COUNT	267 a	354 c	322 g	214 t		
ORIGIN						
Alignment Scores:	4.43e-100	Length:	1157			
Pred. No.:	1024.00	Matches:	193			
Score:	93.72%	Conservative:	16			
Best Local Similarity:	86.55%	Mismatches:	14			
Query Match:	81.79%	Indels:	2			
DB:	14	Gaps:	0			
US-09-830-144-2	COPY_76_303	(1-228)	CD508913	(1-1157)		
Qy	1	ValGluLeuArgGlnLeuSerArgValAsnHisProAsnLeuValLysLeuTyrGlyAla	20			
Db	226	GTGAGCTCCGCCAGCTGTCGCCGGTGAATCACCCCAATATGTGAAGCTGTATGGCTCG	285			
Qy	21	CysLeuAnpProValCysLeuValMetCysLeuAlaGluGlySerLeuTyrAsnVal	40			
Db	286	TGTGACCAACCCAGCTGCTGCTATGAGATGATGAGATGCGGCTCTCTGTATAACCTC	345			
Qy	41	LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu	60			
Db	346	CTGCACAGTGGCGACCCCGCCACCTACACGGCGTCCACGCCCATGAGCTGTGCTG	405			
Qy	61	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArg	80			



directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH MGC Library."

BASE COUNT 253 a 243 c 298 g 266 t 2 others

Alignment Scores:  
Pred. No.: 2,85e-99 Length: 1062  
Score: 1016.00 Matches: 203  
Percent Similarity: 94.52% Conservative: 4  
Best Local Similarity: 92.63% Mismatches: 5  
Query Match: 81.15% Indels: 7  
DB: 12 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x BM554120 (1-1062)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20  
Db GTAGAGCTTCGCGACTTATCCGCTGTAACCATCTTAATATTGTAAGCTTTATGAGCC 467  
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
Db TGCTTGAATCCAGTGCTTGTGTGATGAATATGCTGAAGGGGGCTCTTTATATATG 527  
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
Db CTGCATGTGTGTAACCATTCCTATATATCTGCTGCCACGCAATGAGTTGGTGT 587  
QY 61 GlnCysSerGlnGlyValAlaTyrIleHisSerMetGlnProLysAlaLeuIleHisArg 80  
Db CAGTGTTCCTCCAAAGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTTAATTCACAGG 647  
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
Db GACCTGAACCAACCAACTTACTCTGCTGAGGGGGAGCAGTCTTAAATAATTTGTGAT 707  
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLeuLysGlySerAlaAlaTyr 120  
Db TTGGTACAGCTGTGACATTCAGACACATGACCAATACCAAGGGGGAGTGGCTGTG 767  
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140  
Db ATGGCACTTGAAGTGTGAGGTAGTAAATTACAGTGAATAATGTGAGCTTCAGCTGG 827  
QY 141 GlyIleIleLeuTyrGluValIleThrArgArg-LysProPheAspGluIleGly-P 160  
Db GGTATATATCTTTGGGAAGTGATAACCGCTCGGAAACCCCTTTGATGAGATGGTGGCC 887  
QY 160 roAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeu-IleLysAsn 179  
Db CAGCTTTCGAATCATGTGGCTGGTCAATATGATCTGACCACTGATATAAAAAAT 947  
QY 180 LeuProLysProIleGlu-SerLeuMetThr-ArgCysTyrSerLysAspPro-SerGln 198  
Db TTACCTAGGCCCATTTGAAAGCCTGATGATCTCCNTGNTTGGGTAAAGATCTTTCCAG 1007  
QY 199 ArgProSerMetGlu-GluIleValIysIleMetThrHisLeu 212  
Db CGCCCTTTATGAGGGAATGTTGTAATAATGATCACTGCTGCTG 1050

RESULT 10  
LOCUS BU104161  
DEFINITION 60300594F1 CSEQCHL01 Gallus gallus linear EST 25-NOV-2002  
sequence.  
ACCESSION BU104161  
VERSION BU104161.1  
KEYWORDS EST, GI:25304988  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus  
1 (bases 1 to 750)  
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E., Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.  
A Comprehensive Collection of Chicken cDNAs  
Curr. Biol. 12 (22), 1965-1969 (2002)  
22335534  
12445392  
Contact: Simon Hubbard  
Department of Biomolecular Sciences  
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK  
Tel: 01612008930  
Fax: 01612360409  
Email: Simon.Hubbard@umist.ac.uk.  
Location/Qualifiers  
1..750  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="White Leghorn, Hisex"  
/db\_xref="taxon:9031"  
/clone="CHES123K24"  
/tissue\_type="whole embryo"  
/dev\_stage="20-21"  
/lab\_host="DH10B"  
/clone\_lib="CSEQCHL01"  
/note="Organ: whole embryo; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BspI and BamHI sites [5'ggcgcgtgcagccggatccgaaagaaag] [5'aattcttttggatccgggtgcagc]"

BASE COUNT 222 a 146 c 181 g 201 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.47e-97 Length: 750  
Score: 998.00 Matches: 194  
Percent Similarity: 97.52% Conservative: 3  
Best Local Similarity: 96.04% Mismatches: 4  
Query Match: 79.71% Indels: 3  
DB: 13 Gaps: 1

US-09-830-144-2\_COPY\_76\_303 (1-228) x BU104161 (1-750)  
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20  
Db GTAGAGCTTCGCAACTGTCAAGTAACCATCTTAATATTGTAAGTTATGAGCC 208  
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40  
Db T--CTAAACCCAGTGCTTTGTTATGGATGATGCTGAAGGAGGTTCTCTGTACAATGTG 266  
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60  
Db CTGCATGTGTGCTGAACCTCTGCTCTCATATATCTGTCGACACGCAATGATGTGGTGT 326  
QY 61 GlnCysSerGlnGlyValAlaTyrIleHisSerMetGlnProLysAlaLeuIleHisArg 80  
Db CAGTGTTCCTCCAAAGAGTGGCATCTTCACAGTATGAAACCAAGCCCTAATTCACAGA 386  
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100  
Db GACCTGAACCAACCAAAATTTGCTTCTGTAGTGGGGGACAGTTCTTAAAGATCTGTGAT 446

Thu Dec 4 17:00:24 2003

```

QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 447 TTGTGTCAGCCTGTGATATTCAACACACACATGACCAACAATAAGGGAAGTCTCTGG 506
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db 507 ATGGCACCTGAAGTTTTTGAAGGTAGCAATTACAGTGAAATGATGACGTTTCAAGTTGG 566
QY 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyPro 160
Db 567 GGTATATATCTCTGGAGGTATATACCCGTAGAAACCTTTTGATGAGATCGTGGTCCA 626
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 627 GCTTTCCGCAATAATGTGGCGAGTTCACAAATGTGTACTCCACCACTGATCAAAACTTA 686
QY 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
Db 687 CCTAAACCAATTGAGAG-TTAATGACCCGCTGTGTCCAGATCTCTACA---AGACCT 742
QY 201 SerMet 202
Db 743 TCCATG 748

RESULT 11
BJ073883
LOCUS
DEFINITION
BJ073883 646 bp mRNA linear EST 11-DEC-2001
laevis cDNA clone XL101n05 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 646)
REFERENCE
1 Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
Y.
Expressed genes in X. laevis embryo
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.
Location/Qualifiers
1 . 646
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL101n05"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
BASE COUNT 190 a 132 c 143 g 181 t
ORIGIN

Alignment Scores:
Pred. No.: 3,83e-96 Length: 646
Score: 984.00 Matches: 182
Percent Similarity: 97.3% Conservative: 3
Best Local Similarity: 95.7% Mismatches: 5
Query Match: 78.5% Indels: 1
DB: 12 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BJ073883 (1-646)
QY 39 AsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrp 58

```

```

Db 4 AATGTTTTCATGAGCTGAACCTTTGCTTTACAT-ACTGCTGCCATGCAATGAGTTGG 62
QY 59 CysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeu 78
63 TGTTTACAAATGTGCCAAGGAGTTGCATATTATTACATAGCATGAGCCAAAGGCTCTGATT 122
QY 79 HisArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIle 98
123 CACAGGACCTCAACACCAACCAACTTTGTTGTTGTAGCTGGAGGCACCTGTTCTTAAGATT 182
QY 99 CysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAla 118
183 TGTGACTTTGTTGACAGCCTGTGATATTTCAGACTCATCTACTATAAACAAGAGAGTGA 242
QY 119 AlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPhe 138
243 GCATGGATGTCCTCCAGAAAGTTTTTGAAGGTAGCACTACAGCGAAAATGTGACGTTT 302
QY 139 SerTrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGly 158
303 AGTTGGGCAATTATTCTTTGGGAAGTATAAACCAGAAAGAAACCTTTTCGATGAAATTGGT 362
QY 159 GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLys 178
363 GTGCCAGCCTTCGTATAATGTGGGCTGTTCAATGGTACTGGGCCACCAITTAATAA 422
QY 179 AsnLeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGln 198
423 AATTTGCCCTAAGCCTATTGAAAGCTTAATGACTCGCTGTGTTCCAAAGATCCCCCAA 482
QY 199 ArgProSerMetGluGluIleValIleMetThrHisLeuMetArgTyrPheProGly 218
483 AGACCTTCAATGAGGAGATTGTCAAGATAATCACACATCTAAAGCAGTATTTTCTGTGA 542
QY 219 AlaAspGluProLeuGlnTyrProCysGln 228
543 GCACAGCTTCTCTTACATGATCTCTGTGTCAG 572
Db

RESULT 12
BJ062988
LOCUS
DEFINITION
BJ062988 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL069ml7 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 688)
REFERENCE
1 Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
Y.
Expressed genes in X. laevis embryo
Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.
Location/Qualifiers
1 . 688
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL069ml7"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
FEATURES
source

```

```

BASE COUNT 206 a 139 c 154 g 188 t 1 others
ORIGIN

Alignment Scores:
Score: 6.95e-96 Length: 688
Percent Similarity: 97.83% Matches: 177
Best Local Similarity: 96.20% Conservative: 3
Query Match: 78.43% Mismatches: 4
DB: 12 Indels: 0
Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BJ062988 (1-688)

QY 45 GluProLeuProTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGln 64
Db 22 GAACCTTTGCCCTTACTACTACTGCTGCCCATGCAATGAGTTGGTGTTCATATGTCGCCAA 81
QY 65 GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArgAspLeuLysPro 84
Db 82 GGAGTTGCATATTACATAGCATGAAGCCAAAGGCTCTGATTACAGGACCTCAACCA 141
QY 85 ProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAspPheGlyThrAla 104
Db 142 CCAAACTTGTCTGCTAGTCTGAGGACCTGTTCTTAAGATTGTGACTTTGGTACAGCC 201
QY 105 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGlu 124
Db 202 TGTGATATTTCAGACTCACATGACTTAATCAAGAGGAGTGCAGCATGATGCTCCAGAA 261
QY 125 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleLeu 144
Db 262 GTTTTGAAGTAGCACTACAGCGAAAATGTGACGTGTTAGTTGGGCATTATTCTT 321
QY 145 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 164
Db 322 TGGGAAGTAATAACCCGAGAAACCTTTCGATGAATTTGGTGTCACAGCGTTCGTATA 381
QY 165 MetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeuProLysProIle 184
Db 382 ATGTGGCTGTTCACAAATGCTGCTGCTCAAGATATTAATAAAATTTGCCCTAAGCCATT 441
QY 185 GluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGlu 204
Db 502 ATGTCAAGATATGACACATCTAAAGCAGTATTTCTCTGGAGCAGACGTTTCTCTTACAG 561
QY 225 TyrProCysGln 228
Db 562 TATCCTTGTTCAG 573

RESULT 13
BJ074867
LOCUS
DEFINITION
696 bp mRNA linear EST 11-DEC-2001
laevis cDNA clone XL071110 5', mRNA sequence.
ACCESSION
BJ074867
VERSION
BJ074867.1 GI:17505056
KEYWORDS
EST.
SOURCE
Xenopus laevis (African clawed frog)
ORGANISM
Xenopus laevis
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 696)
REFERENCE
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
,Y.
Expressed genes in X. laevis embryo
Unpublished
Contact: Tadasu Shin-i

TITLE
Unpublished
COMMENT

```

Center For Genetic Resource Information  
National Institute of Genetics  
1111 Yata, Mishima, Shizuoka 411-8540, Japan  
Tel: 81-559-81-6856  
Fax: 81-559-81-6855  
Email: tshini@genes.nig.ac.jp.

#### FEATURES

source  
1..696  
Location/Qualifiers  
/organism="Xenopus laevis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8355"  
/clone="XL071110"  
/tissue\_type="whole embryo"  
/dev\_stage="stage 25"  
/clone\_lib="NIBB Mochii normalized Xenopus tailbud library"

BASE COUNT 210 a 142 c 155 g 188 t 1 others  
ORIGIN

#### Alignment Scores:

Pred. No.: 7.07e-96 Length: 696  
Score: 982.00 Matches: 177  
Percent Similarity: 97.83% Conservative: 3  
Best Local Similarity: 96.20% Mismatches: 4  
Query Match: 78.43% Indels: 0  
DB: 12 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x BJ074867 (1-696)

```

QY 45 GluProLeuProTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGln 64
Db 22 GAACCTTTGCCCTTACTACTGCTGCCCATGCAATGAGTTGGTGTTCATATGTCGCCAA 81
QY 65 GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArgAspLeuLysPro 84
Db 82 GGAGTTGCATATTACATAGCATGAAGCCAAAGGCTCTGATTACAGGACCTCAACCA 141
QY 85 ProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAspPheGlyThrAla 104
Db 142 CCAAACTTGTCTGCTAGTCTGAGGACCTGTTCTTAAGATTGTGACTTTGGTACAGCC 201
QY 105 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGlu 124
Db 202 TGTGATATTTCAGACTCACATGACTTAATCAAGAGGAGTGCAGCATGATGCTCCAGAA 261
QY 125 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleLeu 144
Db 262 GTTTTGAAGTAGCACTACAGCGAAAATGTGACGTGTTAGTTGGGCATTATTCTT 321
QY 145 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 164
Db 322 TGGGAAGTAATAACCCGAGAAACCTTTCGATGAATTTGGTGTCACAGCGTTCGTATA 381
QY 165 MetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeuProLysProIle 184
Db 382 ATGTGGCTGTTCACAAATGCTGCTGCTCAAGATATTAATAAAATTTGCCCTAAGCCATT 441
QY 185 GluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGlu 204
Db 442 GAAAGCTTAATGACTGCTGCTGCTGCTCAAGATATTCCTCCCAAGACCTTCAATGGAGGAG 501
QY 205 IleValLysIleMetThrHisLeuMetArgTyrPheProGlyValAlaAspGluProLeuGln 224
Db 502 ATGTCAAGATATGACACATCTAAAGCAGTATTTCTCTGGAGCAGACGTTTCTCTTACAG 561
QY 225 TyrProCysGln 228
Db 562 TATCCTTGTTCAG 573

```

#### RESULT 14

CB112561

LOCUS

DEFINITION K-EST0154393 L6ChoCKO Homo sapiens cDNA clone L6ChoCKO-4-G09 5',

600 bp mRNA linear EST 28-JAN-2003

61 GlnCysSerGlnGlyValAlaTyrIleuHisSerMetGlnProLysAlaLeuIleHisarg 80  
 251 CAGTGTTCACAGGAGTGGCTTACTTACAGCAGTGAACCAAGCGCTAATTCACAGG 310  
 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysasp 100  
 311 GACCTGAAACACCAAACTTACTGCTGGTTCAGGGGGGACAGTCTTAAAAATTTGTGAT 370  
 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120  
 371 TTGTGACAGCCTGTGACATTCAGACACATGACATGACATGACATGACATGACATG 430  
 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140  
 431 ATGGCACCTGAAGTTTGAAGGTAGTAAATACAGTGAAGAAATGACCGCTTCAGCTG 490  
 141 GlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAsnGluLeuGlyGlyPro 160  
 491 GGTATTATTCTTTGGAGAGTGAACGCTCGAAACCCCTTTGATGAGATTGGTGCCCA 550  
 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeu 176  
 551 GCTTCCGAATCATGTGGGCTGTTTATATGTTACTGACCCACTG 598

RESULT 15  
 BF780358 910 bp mRNA linear EST 12-JAN-2001  
 LOCUS 602103276F1 NCI\_CGAP\_Kid14 Mus musculus cDNA clone IMAGE:4221379  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BF780358  
 VERSION BF780358.1 GI:12085481  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 910)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs@mail.nih.gov  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9806 row: h column: 20  
 High quality sequence stop: 664.  
 Location/Qualifiers  
 1..910  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4221379"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_Kid14"  
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site: 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.  
 Average insert size 1.75 kb. Constructed by Life  
 Technologies. Note: this is a NCI CGAP Library. |"  
 BASE COUNT 247 a 209 c 246 g 207 t  
 ORIGIN

Accession CB112561 GI:27938368  
 VERSION CB112561.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 600)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 4 row: G column: 09  
 High quality sequence stop: 600.  
 Location/Qualifiers  
 1..600  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="L6ChoCK0-4-G09"  
 /sex="M"  
 /cell\_line="Cho-CK"  
 /lab\_host="Top10P"  
 /clone\_lib="L6ChoCK0"  
 /note="Organ: Liver; Vector: pcNS-D2; Site: 1: EcoRI;  
 Site 2: NotI; The poly (A)+ RNA was dephosphorylated with  
 bacterial alkaline phosphatase (BAP) and then decapped  
 with tobacco acid pyrophosphatase (TAP). The decapped  
 intact mRNA was ligated with DNA-RNA linker including  
 EcoRI site by treatment of T4 RNA ligase and the first  
 strand cDNA was synthesized from oligo dT-selected mRNA by  
 priming with dr-tailed vector. The dr-tailed vector was  
 adjusted to have about 60nt. The cDNA vector was  
 circularized with E. coli DNA ligase after digestion of  
 EcoRI which site is also included in vector. An RNA strand  
 converted to a DNA strand by Okayama-Berg method. The  
 obtained cDNA vectors were used for transfection of  
 competent cells E. coli Top10P by electroporation method.  
 The cDNA libraries constructed by this method are  
 full-length enriched cDNA library."  
 BASE COUNT 164 a 113 c 154 g 169 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,036-94 Length: 600  
 Score: 966.00 Matches: 176  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 77.16% Indels: 0  
 DB: 14 Gaps: 0

US-09-830-144-2\_COPY\_76\_303 (1-228) x CB112561 (1-600)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20  
 DB 71 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATGTAAGCTTTATGAGCC 130  
 QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40  
 DB 131 TCGTTGAATCCAGTGTCTGTGTGTAATATGCTGAAGGGGGCTCTTTATATATG 190  
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60  
 DB 191 CTGATGGGTGCTGAACCAITGGCCATATTATATCTGCTGCCACCATGATGTTGGTTT 250

```
Query Match: 76.16% Indels: 12
DB: 10 Gaps: 1
US-09-830-144-2_COPY_76_303 (1-228) x BF780358 (1-910)

QY 28 ValMetGluTyrAlaGluGlySer-Leu-----TyrAsnVa 40
Db 4 GTGATGGAAATATGCAGAGGGGGCTCATGTATAATGTTTGTGCTTCTTCTTACACT 63
QY 40 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTipCysIle 60
Db 64 GCTGCATGCTGCTGAACCATTCCTTACTACTGCTGCTCATGCCATGAGCTGGTGT 123
QY 60 uClnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisAr 80
Db 124 ACAGTGTTCCTCAAGAGGTGGCTTACTGTCACAGCATGCAGCCCAAGCCGCTGATTCACAG 183
QY 80 GAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAs 100
Db 184 GGACCTCAAGCCTCCAAACTTCTGCTGTTGAGGAGGGACAGTTCTAAAAATCTGGCA 243
QY 100 pPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTr 120
Db 244 TTTGGTACAGCTTGTCATCCCAACACACATGACCAATATAAAGGAGTGCTGTG 303
QY 120 pMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTr 140
Db 304 GATGGCGCTGAAGTGTGTTGAAGGTAGCAATTACAGTGAAGAAGTGTGATGCTTCAGCTG 363
QY 140 pGlyIleLeuTyrGluValIle-ThrArgArgLysProPheAspGluIleGlyGlyP 160
Db 364 GGGTATTATCTCTGGGAAGTGATAAACACGCGGAAACCTTCGATCAGATCGGTGGCC 423
QY 160 roAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnL 180
Db 424 CAGCTTTCAGATCATGTGGGCTGTTTATAATGGCCTCGACCCACCACTGATCAAAAAT 483
QY 180 euProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgp 200
Db 484 TACCTAAGCCCATTTGAGAGCTTGATGACACGCTG-TGGTCTAAGGACCCCATCTCAGCG-C 541
QY 200 roSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaA 220
Db 542 CTTCAATGGAGGAATGTGAAAATAATGACTCACTTGTATGCGGTACTTCCAGGAGGG 601
QY 220 spGluProLeuGlnTyrProCysGln 228
Db 602 ATGAGCA-TTACAGTATACTTGTCAG 626
```

Search completed: December 4, 2003, 12:42:49  
Job time : 2906.3 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 4, 2003, 09:09:50 ; Search time 22.3378 Seconds  
(without alignments)  
479.997 Million cell updates/sec

Title: US-09-830-144-2\_COPY\_76\_303  
Perfect score: 1252  
Sequence: 1 VELQLSRVNHPIVLYGA.....MTHLMRYFPGADPLQYPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 segs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1252	100.0	579	1	M3K7_MOUSE	Q62073 mus musculus
2	1252	100.0	606	1	M3K7_HUMAN	O43318 homo sapien
3	435.5	34.8	393	1	M3K7_DROME	P83104 drosophila
4	325.5	31.3	954	1	M3K4_HUMAN	O27779 homo sapien
5	363.5	29.0	394	1	M3K9_HUMAN	P80192 homo sapien
6	355.5	28.4	888	1	M3K3_RAT	Q63796 rattus norv
7	353.5	28.2	859	1	M3K3_HUMAN	Q2852 homo sapien
8	353.5	28.2	888	1	M3K3_MOUSE	Q60700 mus musculus
9	341	27.2	410	1	KTK2_DICDI	P18161 dictyosteli
10	323	25.8	821	1	CTR1_ARATH	Q05609 arabidopsis
11	307.5	24.6	1115	1	RET_MOUSE	P35546 mus musculus
12	304.5	24.3	1114	1	RET_HUMAN	P07943 homo sapien
13	301.5	24.1	630	1	TEC_MOUSE	P24604 mus musculus
14	301.5	24.1	631	1	TEC_HUMAN	P42680 homo sapien
15	299.5	23.9	625	1	ITK_MOUSE	Q03526 mus musculus
16	298.5	23.8	620	1	ITK_HUMAN	Q08881 homo sapien
17	294.5	23.5	2347	1	KROS_HUMAN	P08922 homo sapien
18	292.5	23.4	402	1	KROS_AVISU	P00529 avian sarco
19	292	23.3	507	1	KROS_CHICK	P08941 gallus gall
20	290	23.2	1584	1	KTK2_DICDI	P18160 dictyosteli
21	288.5	23.0	628	1	TES1_RAT	Q63572 rattus norv
22	288	23.0	1307	1	PHYL_CERPU	P25848 ceratodon p
23	286	22.8	804	1	FPS_DROME	P18106 drosophila
24	284.5	22.7	832	1	ANR3_HUMAN	P57078 homo sapien
25	283.5	22.6	527	1	TXK_MOUSE	P42682 mus musculus
26	280	22.4	822	1	FGRI_RAT	Q04589 rattus norv
27	279	22.3	467	1	MATK_RAT	P41243 rattus norv
28	278	22.2	822	1	FGRI_HUMAN	P11362 homo sapien
29	278	22.2	822	1	FGRI_MOUSE	P16092 mus musculus
30	277.5	22.2	786	1	BTCL_DROME	P08630 drosophila
31	275	22.0	984	1	EPK3_RAT	Q08680 rattus norv
32	274.5	21.9	801	1	FGRI_MOUSE	Q61851 mus musculus
33	271.5	21.7	1062	1	CC7_SCHPO	P41892 schizosacch

P04627 mus musculus  
P14056 rattus norv  
P50545 rattus norv  
P22607 homo sapien  
P42679 homo sapien  
Q15569 homo sapien  
P00541 avian sarco  
P21804 gallus gall  
P29320 homo sapien  
P54736 myxococcus  
Q05397 homo sapien  
O35346 rattus norv

34 271 21.6 604 1 KRAA\_MOUSE  
35 271 21.6 604 1 KRAA\_RAT  
36 270.5 21.6 503 1 HCK\_RAT  
37 270 21.6 806 1 FGR3\_HUMAN  
38 269.5 21.5 507 1 MATK\_HUMAN  
39 269.5 21.5 626 1 TES1\_HUMAN  
40 269 21.5 533 1 FPS\_AVISP  
41 269 21.5 819 1 FGR1\_CHICK  
42 269 21.5 983 1 EPA3\_HUMAN  
43 268 21.4 830 1 PKN2\_MYXXA  
44 268 21.4 1052 1 FAK1\_HUMAN  
45 268 21.4 1055 1 FAK1\_RAT

## ALIGNMENTS

RESULT 1  
M3K7\_MOUSE  
ID M3K7\_MOUSE STANDARD; PRT; 579 AA.  
AC Q62073;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.1-)  
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).  
GN MAP3K7 OR TAK1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96132377; PubMed=8533096;  
RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,  
RT "Identification of a member of the MAPKKK family as a potential mediator of TGF-beta signal transduction."  
RL Science 270:2008-2011(1995).  
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.  
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B ACTIVATION.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
CC MAP KINASE KINASE KINASE SUBFAMILY.  
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D76446; BAA1184.1; --  
CC HSPSP; P08631; IAD5.  
CC MGD; MGI:1346877; Map3k7.  
CC InterPro; IPR000719; Prot\_kinase.  
CC InterPro; IPR002290; Ser\_Thr\_pkinase.  
CC InterPro; IPR001245; Tyr\_pkinase.  
CC Pfam; PF00069; pkinase; 1.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
CC PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
CC Transferase; Serine/threonine-protein kinase; ATP-binding.  
CC DOMAIN 8 16 POLY-SER.  
CC FT DOMAIN 36 291  
CC FT NP\_BIND 42 50  
CC FT BINDING 63 63  
CC FT ACT\_SITE 156  
CC FT BY SIMILARITY.  
CC FT BY SIMILARITY.

SQ	SEQUENCE	579 AA; 64227 MW; 97C8F6F3C8E2B3BE CRC64;	
	Query Match	100.0%; Score 1252; DB 1; Length 579;	
	Best Local Similarity	100.0%; Pred. No. 2.5e-110;	
	Matches 228; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	VELRQLSRVNHNPVIVKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 60	
Db	76	VELRQLSRVNHNPVIVKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 135	
QY	61	QCSQGVAYLHSMOPKALIHRLDLPNNLLVAGTGLKICDFGTACDIQTHMTNKGSAAW 120	
Db	136	QCSQGVAYLHSMOPKALIHRLDLPNNLLVAGTGLKICDFGTACDIQTHMTNKGSAAW 195	
QY	121	MAPEVEFGSNYSKCDVFSWGIIWVITRRKFDGIGGPAFRIMWAVHNGTRPPLIKNL 180	
Db	196	MAPEVEFGSNYSKCDVFSWGIIWVITRRKFDGIGGPAFRIMWAVHNGTRPPLIKNL 255	
QY	181	PKPIESLMTRCWSKDPQRSMEBEIVKIMTHLMRYFFGADEPLQYPCQ 228	
Db	256	PKPIESLMTRCWSKDPQRSMEBEIVKIMTHLMRYFFGADEPLQYPCQ 303	
RESULT 2			
M3K7	HUMAN	STANDARD; PRT; 606 AA.	
ID	M3K7 DROME		
AC	O43318; O43317; O43319;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.1-)		
DE	(transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).		
GN	MAP3K7 OR TAK1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Carnivora; Insectivora; Hominoidea; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1A; 1B AND 1C).		
RC	TISSUE=Lung;		
RC	MEDLINE=98153801; PubMed=9480845;		
RA	Sakurai H., Shigemori N., Hasegawa K., Sugita T.;		
RT	"TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism.";		
RL	Biochem. Biophys. Res. Commun. 243:545-549(1998).		
CC	1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.		
CC	1- MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B ACTIVATION.		
CC	1- ALTERNATIVE PRODUCTS:		
CC	Event=Alternative splicing; Named isoforms=3;		
CC	Name=1B;		
CC	Isoid=O43318-1; Sequence=Displayed;		
CC	Name=1A;		
CC	Isoid=O43318-2; Sequence=VSP_004886;		
CC	Name=1C;		
CC	Isoid=O43318-3; Sequence=VSP_004887, VSP_004888;		
CC	1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP KINASE KINASE KINASE SUBFAMILY.		
CC	1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	EMBL; AB009357; BAA25026.1; -		
DR	EMBL; AB009356; BAA25025.1; -		
DR	EMBL; AB009358; BAA25027.2; -		
DR	PIR; JC5955; JC5955.		
DR	PIR; JC5956; JC5956.		
DR	HSP; P08631; LAD5.		
DR	Genew; HGNC:6859; MAP3K7.		
DR	MIM; 602614; -		
DR	GO; GO:0004709; F:MAP kinase kinase kinase activity; TAS.		
DR	GO; GO:0007179; P:TGFbeta receptor signaling pathway; TAS.		
DR	InterPro; IPR000719; Prot kinase.		
DR	InterPro; IPR002290; Ser Thr kinase.		
DR	InterPro; IPR001245; Tyr_kinase.		
DR	Pfam; PF00069; pkinase; 1.		
DR	PRINTS; PR00109; TYRKINASE.		
DR	ProDom; PD000001; Prot kinase; 1.		
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.		
DR	PROSITE; PS00108; PROTEIN KINASE ST; 1.		
DR	PROSITE; PS00111; PROTEIN KINASE_DOM; 1.		
KW	Transferase; Serine/threonine-protein kinase; ATP-binding; Alternative splicing.		
KW	DOMAIN 8 14		
FT	POLY-SER.		
FT	DOMAIN 36 291		
FT	DOMAIN 42 50		
FT	NP_BIND 63 63		
FT	BINDING 156 156		
FT	ACT_SITE 404 430		
FT	VARSPPLIC 509 518		
FT	FTID=VSP_004886.		
FT	PLAPCPNSKE -> ARTSCTRTGPG (in isoform 1C).		
FT	FTID=VSP_004887.		
FT	Missing (in isoform 1C).		
FT	FTID=VSP_004888.		
FT	/FTID=VSP_004886.		
SQ	SEQUENCE 606 AA; 67196 MW; 3D8F8147CD174013 CRC64;		
	Query Match 100.0%; Score 1252; DB 1; Length 606;		
	Best Local Similarity 100.0%; Pred. No. 2.6e-110;		
	Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	VELRQLSRVNHNPVIVKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 60	
Db	76	VELRQLSRVNHNPVIVKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 135	
QY	61	QCSQGVAYLHSMOPKALIHRLDLPNNLLVAGTGLKICDFGTACDIQTHMTNKGSAAW 120	
Db	136	QCSQGVAYLHSMOPKALIHRLDLPNNLLVAGTGLKICDFGTACDIQTHMTNKGSAAW 195	
QY	121	MAPEVEFGSNYSKCDVFSWGIIWVITRRKFDGIGGPAFRIMWAVHNGTRPPLIKNL 180	
Db	196	MAPEVEFGSNYSKCDVFSWGIIWVITRRKFDGIGGPAFRIMWAVHNGTRPPLIKNL 255	
QY	181	PKPIESLMTRCWSKDPQRSMEBEIVKIMTHLMRYFFGADEPLQYPCQ 228	
Db	256	PKPIESLMTRCWSKDPQRSMEBEIVKIMTHLMRYFFGADEPLQYPCQ 303	
RESULT 3			
M3K7	DROME	STANDARD; PRT; 393 AA.	
ID	M3K7 DROME		
AC	P83104;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Putative mitogen-activated protein kinase kinase 7 (EC 2.7.1.1-)		
GN	TAK1L.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Berkeley;		
RC	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		



QY	2	ELRQLSRVNHPIVKLYGACLM--PVCVLVMEYAEAGSSLYNVHLHGABPLPYTYTAAHMSWC	59
DB	54	EITHLSSEIDHENVIRVIGRASNGKKDYLLMEYLEEGSLHNYLYGDDKWE-YTVEQAVRMA	112
QY	60	LOCSGOAYLHSMQPKALIHRLDKLPNLLVAGGTIVLKICDGTACDIOTHMTNNKGSAA	119
DB	113	LQCAKALAYLHSLD-RPIVHRDIKPNMLLYNQHEDLKICDFGLATDMSNNKTDMQGTLR	171
QY	120	WMAPEVFGSGNSYSEKCDVFSWGIIILWEVITRRKPEDEIGGP--APRIMWAVHNGTAPPL-	176
DB	172	YMAPEAIKHLAYTAKCDVYSGIIMLWELWTRQLPYSHLENPNYSQVAIMKAISSGEKLPME	231
QY	177	--IKNLKPIESLWTRCWSKDPSPQSPSMEEIVKIMTHLMRYFPFGADEPIQYP	226
DB	232	AVRSDCPGIGIKQLMECCMDINPEKRPMSKEIEKFLGE--QYESGTDDEDFIKP	281
RESULT 4			
M3KA_HUMAN			
ID	M3KA_HUMAN	STANDARD; PRT; 954 AA.	
AC	002779; Q12761; Q14871;		
DT	01-FEB-1994 (Rel. 28, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Mitogen-activated protein kinase kinase kinase 10 (EC 2.7.1.37)		
DE	(Mixed lineage kinase 2) (protein kinase MST).		
GN	MAP3K10 OR MLK2 OR MST.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=96128179; PubMed=8536694;		
RA	Dorow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,		
RA	Sutherland G.R., Simpson R.J.;		
RT	"Complete nucleotide sequence, expression, and chromosomal		
RT	localisation of human mixed-lineage kinase 2.";		
RL	Eur. J. Biochem. 234:492-500(1995).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=95249256; PubMed=7731697;		
RA	Katoh M., Hirai M., Sugimura T., Terada M.;		
RT	"Cloning and characterization of MST, a novel (putative)		
RT	serine/threonine kinase with SH3 domain.";		
RL	Oncogene 10:1447-1451(1995).		
RN	[3]		
RP	SEQUENCE OF 244-480 FROM N.A.		
RC	TISSUE=colon epithelium;		
RX	MEDLINE=93238756; PubMed=8477742;		
RA	Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;		
RT	"Identification of a new family of human epithelial protein kinases		
RT	containing two leucine/isoleucine-zipper domains.";		
RL	Eur. J. Biochem. 213:701-710(1993).		
CC	-1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.		
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.		
CC	-1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.		
CC	MAP KINASE KINASE SUBFAMILY.		
CC	-1- SIMILARITY: Contains 1 SH3 domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
EMBL	X90846; CAA62351.1; -		
DR	EMBL; Z48615; CAA88531.1; -		



```

DE (MAPK-upstream kinase) (MUK).
GN MAP3K12 OR MJK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euarchia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_taxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96226099; PubMed=8637721;
RA Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.;
RT "Activation of the JNK pathway by distantly related protein kinases,
RT MEKK and MUK.";
RL Oncogene 12:641-650(1996).
CC -! FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -! CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -! COFACTOR: Magnesium.
CC -! SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -! PMW: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -! SIMILARITY: BELONGS TO THE SER/TRH FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; D49785; BAA08621.1; -.
DR HSSP; P12931; IFMK.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00089; pkinase_1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR Transfaser; Serine/threonine-protein kinase; ATP-binding;
KW phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399 PROTEIN_KINASE.
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185 ATP (BY SIMILARITY).
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62 POLY-GLY.
FT DOMAIN 668 671 POLY-PRO.
FT DOMAIN 698 701 POLY-PRO.
FT DOMAIN 753 758 POLY-GLU.
SQ SEQUENCE 888 AA; 96307 MW; 52AD964006ABE149 CRC64;

Query Match 28.4%; Score 355.5; DB 1; Length 888;
Best Local Similarity 34.7%; Pred.No. 1.le-25;
Matches 75; Conservative 44; Mismatches 80; Indels 17; Gaps 8;

QY 2 ELRLSRVNHNPITVKLYGACLNP--VCVLNVEAEGGSLLNVHLGAEPFLPYTTAAAHMSWC 59
Db :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
194 DIKLAKLKHPNIITFGVCTQAPCYCILMEFCACQLFEVLGRPV---TFSLLDVWS 250
QY 60 LQSQQVAYLHSNQPKALIHRLDKPNLLLVAGGTIVLKICDFGTACDIQTHTNNK--GS 117
Db ::||::||::||::||::||::||::||::||::||::||::||::||::||:
251 MGTAGGMNYLIHLHK----IIHRDLKSPN-MLTIVDDVVKIISDFTSKLSDKSTKMSFACT 306
QY 118 AAWMAPEVFEGSNYSFKCDVSFWSGIILLWEVIIRTKPDFDIGGFAPRMKAV-HNGTRDPL 176
Db :|::||::||::||::||::||::||::||::||::||::||::||::||:
307 VAMMAPEVTNRFPVSEKVDIMISFGVVLBELLGEIPYKDWDSSA--IIMGVGNSIHLPV 364

```

2



DR ProDom: PD000001; Prot kinase; 1.  
DR SMART; SM02220; S\_TK; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
KW Transferase; Serine/threonine-protein kinase; ATP-binding.  
FT DOMAIN 65 69 POLY-GLY.  
FT DOMAIN 135 141 POLY-GLY.  
FT DOMAIN 551 809 PROTEIN KINASE.  
FT NP BIND 557 565 ATP (BY SIMILARITY).  
FT BINDING 578 578 ATP (BY SIMILARITY).  
FT ACT SITE 676 676 BY SIMILARITY.  
FT MUTAGEN 596 596 E-K; IN CTIR1-4; EXHIBITS ETHYLENE-TREATED PHENOTYPE.  
FT MUTAGEN 694 694 D-S; IN CTIR1-1; EXHIBITS ETHYLENE-TREATED PHENOTYPE.  
FT FT  
SQ SEQUENCE 821 AA; 90306 MW; 29223DCDCC15BC CRC64;  
Query Match 25.8%; Score 323; DB 1; Length 821;  
Best Local Similarity 33.9%; Pred. No. 1.2e-22;  
Matches 74; Conservative 45; Mismatches 91; Indels 8; Gaps 6;  
QY 2 ELRQLSRVHNPIVLYGACLP--VCLVMEYAEGLSVNVLHGABLPVYTAHAWSC 59  
Db 596 EVAIMKELRHPIVLFMGAVTQPNLSIVTEYLSRGLYRLHLKSGAREQLDERRRLSMA 655  
QY 60 LQSQGVAYLHMQPKALIHRLDKPNNLLVAGTVLKIDFG-TACDIOTHTNNK--G 116  
Db 656 YDVAKGMYLHNRNP-PIVHRDLKSPNLLVDKRYTV-KVCDPGLSLKASTFLSSKSAAG 713  
QY 117 SAAMVAPEVSGNSYSKCVFSGIILWEVITRRKPFDEIGGPAFRIMWAVNGTRPPL 176  
Db 714 TPWMAFEVURDEPSNEKSIVYSGVGLNRELATLQDPWGNL-NPAQVVAAGVGFCKLEI 772  
QY 177 IKNLKPESILMTRCSKDFSPSEMEIIVKIMHLMR 214  
Db 773 PRNLNPQVAALIEGCWINEPWKPSFATINDLLRPLIK 810  
RESULT 11  
RET MOUSE STANDARD; PRT; 1115 AA.  
AC P35546;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Proto-oncogene tyrosine-protein kinase receptor ret precursor  
DE (EC 2.7.1.112) (C-ret).  
GN RET.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP MEDLINE=930205390; PubMed=8455936;  
RX Iwamoto T., Iatiguchi M., Asai N., Ohkusu K., Nakashima I.,  
RA Takahashi M.;  
RA "cDNA cloning of mouse ret proto-oncogene and its sequence similarity  
RT to the cadherin superfamily.";  
RL Oncogene 8:1087-1091(1993).  
EN [2]  
RP INTERACTION WITH DOK2; DOK4 AND DOK5, PHOSPHORYLATION, AND MUTAGENESIS  
RP OF TYR-1063.  
RX MEDLINE=211363571; PubMed=11470823;  
RA Grimm J., Sachs M., Britsch S., Di Cesare S., Schwarz-Romond T.,  
RA Alitalo K., Birchmeier W.;  
RT "Novel p62dok family members, dok-4 and dok-5, are substrates of the  
RT c-ret receptor tyrosine kinase and mediate neuronal  
RT differentiation.";  
RL J. Cell Biol. 154:345-354(2001).  
CC -!- FUNCTION: Probable receptor with tyrosine-protein kinase activity;  
important for development.

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
tyrosine phosphate.  
CC -!- SUBUNIT: Phosphorylated form interacts with the PBT domain of  
DOK2, DOK4 and DOK5.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Expressed in peripheral nerve cells and  
hematopoietic cells.  
CC -!- PTM: Phosphorylated.  
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.  
CC -!- SIMILARITY: Contains 1 cadherin domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X67812; CAA48013.1; -.  
CC PIR; I48735; S29926.  
CC HSP; P11362; 1FGI.  
CC MGD; MGI:97902; Ret.  
CC GO; GO:0001657; P:retic bud development; IMP.  
CC InterPro; IPR002126; Cadherin.  
CC InterPro; IPR000719; Prot kinase.  
CC InterPro; IPR001245; Tyr.pkinase.  
CC Pfam; PF00028; cadherin; 1.  
CC Pfam; PF00069; pkinase; 1.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom; PD000001; Prot kinase; 1.  
CC SMART; SM00112; CA; 1.  
CC SMART; SM00219; TyKC; 1.  
CC PROSITE; PS0268; CADHERIN 2; 1.  
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
CC PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
CC Transferase; Tyrosine-protein kinase; ATP-binding; Proto-oncogene;  
KW Transmembrane; Phosphorylation; Glycoprotein; Signal.  
FT SIGNAL 1 28 POTENTIAL.  
FT CHAIN 29 1115 PROTO-ONCOGENE TYROSINE-PROTEIN KINASE  
RECEPTOR RET.  
FT FT EXTRACELLULAR (POTENTIAL).  
FT FT POTENTIAL.  
FT FT CYTOPLASMIC (POTENTIAL).  
FT FT CADHERIN.  
FT FT PROTEIN KINASE.  
FT FT ATP (BY SIMILARITY).  
FT FT BINDING 759 759 ATP (BY SIMILARITY).  
FT FT ACT SITE 875 875 BY SIMILARITY.  
FT FT MOD RES 1063 1063 PHOSPHORYLATION.  
FT FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 470 470 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT MUTAGEN 1063 1063 Y->F: ABOLISHES INTERACTION WITH DOK  
PROTEINS.  
SQ SEQUENCE 1115 AA; 123728 MW; 4D75576095C7D2C8 CRC64;  
Query Match 24.6%; Score 307.5; DB 1; Length 1115;  
Best Local Similarity 32.1%; Pred. No. 4.8e-21;  
Matches 79; Conservative 44; Mismatches 78; Indels 45; Gaps 9;  
QY 2 ELRQLSRVHNPIVLYGACLP--VCLVMEYAEGLSVNVLHGABLPVYTAHAWSC 59  
CC -!- FUNCTION: Probable receptor with tyrosine-protein kinase activity;  
important for development.

Db 776 BFNLLKQVNHVHVIKYGACSGDPLLLIIVEYAKYSLRGLRDLRDKIGPAYVGGSGSRN 835  
 QY 51 -----TAAHMSWCLQCSQVAVLHSMQPKALHRLDKPNNLLVAGGTVLKI 98  
 Db 836 SSSLDHPDERVLTGDLISFAWQISGMQVLAEMK---LVHRLAARN-ILVAEGRKMKI 891  
 QY 99 CDFGTACDI---OTHMNNKG--SAAWMAPEVEFGNSYSEKCDVFSWGIILWEVITRRKP 153  
 Db 892 SDFGLSRDVEEDSYVKKSGRIIPVXWMAIESLFDHIYTTQSDVWGFVLLWEIVT---- 947  
 QY 154 FDEIGG-----PAPRIMVAVHNGTTPPLIKNLPKPIESLMTFCWKSQSPRSMEEIVK 207  
 Db 948 ---LGNVPYGPPIPERFLNLLTGHMRPDCSEMYRLMLQCWKQEPDKRPVFADISK 1004  
 QY 208 IHTLM 213  
 Db 1005 DLEKWM 1010  
 RESULT 12  
 ID RET HUMAN STANDARD; PRT; 1114 AA.  
 AC P07349;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DE 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Proto-oncogene tyrosine-protein kinase receptor ret precursor  
 DE (EC 2.7.1.112) (C-ret).  
 GN RET.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 1-280 FROM N.A.  
 RX MEDLINE=89282215; PubMed=2660074;  
 RA Takahashi M.;  
 RT "Isolation of ret proto-oncogene cDNA with an amino-terminal signal  
 RT sequence.";  
 RL Oncogene 4:805-806(1989).  
 RN [2]  
 RP SEQUENCE OF 255-1114 FROM N.A.  
 RX MEDLINE=90272230; PubMed=3078962;  
 RA Takahashi M., Buma Y., Iwamoto T., Inaguma Y., Ikeda H., Hiai H.;  
 RT "Cloning and expression of the ret proto-oncogene encoding a tyrosine  
 RT kinase with two potential transmembrane domains.";  
 RL Oncogene 3:571-578(1988).  
 RN [3]  
 RP SEQUENCE OF 588-1063 FROM N.A.  
 RX MEDLINE=87257826; PubMed=3037315;  
 RA Takahashi M., Cooper G.M.;  
 RT "ret transforming gene encodes a fusion protein homologous to  
 RT tyrosine kinases.";  
 RL Mol. Cell. Biol. 7:1378-1385(1987).  
 RN [4]  
 RP TROSYNE AUTOPHOSPHORYLATION, AND MUTAGENESIS OF TYR-1015 AND  
 RP TYR-1062.  
 RX MEDLINE=20513733; PubMed=11061555;  
 RA Salvatore D., Barone M.V., Salvatore G., Melillo R.M., Chiappetta G.,  
 RA Mineo A., Fenzi G., Vecchio G., Fusco A., Santoro M.;  
 RT "Tyrosines 1015 and 1062 are in vivo autophosphorylation sites in ret  
 RT and ret-derived oncoproteins.";  
 RL J. Clin. Endocrinol. Metab. 85:3898-3907(2000).  
 RN [5]  
 RP REVIEW ON HSCR VARIANTS  
 RX MEDLINE=98023959; PubMed=9359036;  
 RA Hofstra R.M.W., Osinga J., Buys C.H.C.M.;  
 RT "Mutations in Hirschsprung disease: when does a mutation contribute to  
 RT the phenotype.";  
 RL Eur. J. Hum. Genet. 5:180-185(1997).  
 RN [6]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=97220587; PubMed=9067749;  
 RA Mulligan L.M., Eng C., Attie T., Lyonnet S., Marsh D.J., Hyland V.J.,

Eng C., Mulligan L.M.;  
 RT "Mutations of the RET proto-oncogene in the multiple endocrine  
 RT neoplasia type 2 syndromes, related sporadic tumours, and  
 RT Hirschsprung disease.";  
 RL Hum. Mutat. 9:97-109(1997).  
 RN [7]  
 RP VARIANTS MEN2A/FMTC TRP-611; SER-618; ARG-620; TYR-620 AND ARG-634.  
 RX MEDLINE=93372843; PubMed=8103403;  
 RA Donis-Keller H., Dou S., Chi D., Carlson K.M., Toshima K.,  
 RA Laitmore T.C., Howe J.R., Moley J.F., Goodfellow P., Wells S.A. Jr.;  
 RT "Mutations in the RET proto-oncogene are associated with MEN 2A and  
 RT FMTC.";  
 RL Hum. Mol. Genet. 2:851-856(1993).  
 RN [8]  
 RP VARIANTS MEN2A GLY-618; 632-ASP-VAL-ARG-634; GLY-634; PHE-634; TYR-634  
 RP AND SER-634.  
 RX MEDLINE=933725414; PubMed=8099202;  
 RA Mulligan L.M., Kwok J.B.J., Healey C.S., Elsdon M.J., Eng C.,  
 RA Gardner E., Love D.R., Mole S.E., Moore J.K., Papi L., Ponder M.A.,  
 RA Telenius H., Tunncliffe A., Ponder B.A.J.;  
 RT "Germ-line mutations of the RET proto-oncogene in multiple endocrine  
 RT neoplasia type 2A.";  
 RL Nature 363:458-460(1993).  
 RN [9]  
 RP VARIANTS HSCR PRO-40; LEU-399; GLN-762; PRO-765; GLN-897; GLY-972 AND  
 RP LEU-973.  
 RX MEDLINE=95219414; PubMed=7704557;  
 RA Yin L., Barone V., Seri M., Bolino A., Boccardi R., Ceccherini I.,  
 RA Pasini B., Tocco T., Lerone M., Cywes S., Moore S.,  
 RA Vandervinden J.-M., Abramowicz M.J., Kristoffersson U., Larsson L.T.,  
 RA Hamel B.C.J., Silengo M., Martucciello G., Romeo G.;  
 RT "Heterogeneity and low detection rate of RET mutations in Hirschsprung  
 RT disease.";  
 RL Eur. J. Hum. Genet. 2:272-280(1994).  
 RN [10]  
 RP VARIANT MEN2B THR-918.  
 RX MEDLINE=94272459; PubMed=7911697;  
 RA Eng C., Smith D.P., Mulligan L.M., Nagai M.A., Healey C.S.,  
 RA Ponder B.A.J., Gardner E., Scheumann G.F., Jackson C.E., Tunncliffe A.,  
 RA Ponder B.A.J.;  
 RT "Point mutation within the tyrosine kinase domain of the RET  
 RT proto-oncogene in multiple endocrine neoplasia type 2B and related  
 RT sporadic tumours.";  
 RL Hum. Mol. Genet. 3:237-241(1994).  
 RN [11]  
 RP VARIANTS MEN2A/FMTC ARG-618; SER-618; PHE-620; ARG-620; PHE-634;  
 RP GLY-634 AND TYR-634.  
 RX MEDLINE=94348513; PubMed=7915165;  
 RA Xue F., Yu H., Maurer L.H., Memoli V.A., Nutille-Mcmenemy N.,  
 RA Schuster M.K., Browden D.W., Mao J.-I., Noll W.W.;  
 RT "Germline RET mutations in MEN 2A and FMTC and their detection by  
 RT simple DNA diagnostic tests.";  
 RL Hum. Mol. Genet. 3:635-638(1994).  
 RN [12]  
 RP VARIANTS FMTC/MEN2A TYR-609; ARG-618; SER-618 AND SER-620.  
 RX MEDLINE=95152521; PubMed=7849720;  
 RA Blaugrund J.E., Johns M.M. Jr., Eby Y.J., Ball D.W., Baylin S.B.,  
 RA Hruban R.H., Sidransky D.;  
 RT "RET proto-oncogene mutations in inherited and sporadic medullary  
 RT thyroid cancer.";  
 RL Hum. Mol. Genet. 3:1895-1897(1994).  
 RN [13]  
 RP VARIANTS FMTC, AND VARIANTS MEN2A.  
 RX MEDLINE=95179108; PubMed=7874109;  
 RA Schuffenecker I., Billand M., Calender A., Chambe B., Ginot N.,  
 RA Calmettes C., Modigliani E., Lenoir G.M.;  
 RT "RET proto-oncogene mutations in French MEN 2A and FMTC families.";  
 RL Hum. Mol. Genet. 3:1939-1943(1994).  
 RN [14]  
 RP VARIANT HSCR TRP-609, VARIANT HSCR/MEN2A ARG-618, AND VARIANT  
 RP HSCR/FMTC ARG-620.  
 RX MEDLINE=95187155; PubMed=7881414;  
 RA Mulligan L.M., Eng C., Attie T., Lyonnet S., Marsh D.J., Hyland V.J.,

RA Robinson B.G., Frilling A., Verellen-Dumoulin C., Safar A.,  
 RA Venter D.J., Munnich A., Ponder B.A.J.;  
 RT "Diverse phenotypes associated with exon 10 mutations of the RET  
 RT proto-oncogene.";  
 RL Hum. Mol. Genet. 3:2163-2167(1994).  
 RN [15]  
 RP VARIANT MEN2B THR-918.  
 RX MEDLINE=94159102; PubMed=7906866;  
 RA Hofstra R.M.W., Landsvater R.M., Ceccherini I., Stulp R.P.,  
 RA Stefagen T., Luo Y., Pasini B., Hoepfner J.W.M., van Amstel H.K.P.,  
 RA Romeo G., Lips C.J.M., Buys C.H.C.M.;  
 RT "A mutation in the RET proto-oncogene associated with multiple  
 RT endocrine neoplasia type 2B and sporadic medullary thyroid  
 RT carcinoma.";  
 RL Nature 367:375-376(1994).  
 RN [16]  
 RP VARIANTS HSCR PRO-765; GLN-897 AND GLY-972.  
 RX MEDLINE=94159103; PubMed=8114938;  
 RA Romeo G., Ronchetto P., Luo Y., Barone V., Seri M., Ceccherini I.,  
 RA Pasini B., Bocciaardi R., Lerone M., Kaariainen H., Martucciello G.;  
 RT "Point mutations affecting the tyrosine kinase domain of the RET  
 RT proto-oncogene in Hirschsprung's disease.";  
 RL Nature 367:377-378(1994).  
 RN [17]  
 RP VARIANTS HSCR LEU-32; LEU-64; GLN-330 AND LEU-393.  
 RX MEDLINE=94159104; PubMed=8114939;  
 RA Ederly P., Lyonnet S., Mulligan L.M., Pelet A., Dow E., Abel L.,  
 RA Holder S., Nihoul-Fekete C., Ponder B.A.J., Munnich A.;  
 RT "Mutations of the RET proto-oncogene in Hirschsprung's disease.";  
 RL Nature 367:378-380(1994).  
 RN [18]  
 RP VARIANT MEN2B THR-918.  
 RX MEDLINE=94151373; PubMed=7906417;  
 RA Carlson K.M., Dou S., Chi D., Scavarda N., Toshima K., Jackson C.E.,  
 RA Wells S.A., Jr., Goodfellow P.J., Donis-Keller H.;  
 RT "Single missense mutation in the tyrosine kinase catalytic domain of  
 RT the RET protooncogene is associated with multiple endocrine neoplasia  
 RT type 2B.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1579-1583(1994).  
 RN [19]  
 RP VARIANTS MTC: FVTC; MEN2A AND MEN2B.  
 RX MEDLINE=96223053; PubMed=8625130;  
 RA Komminoth P., Kunz E.K., Matias-Guiu X., Hiort O., Christiansen G.,  
 RA Colomer A., Roth J., Heitz P.U.;  
 RT "Analysis of RET protooncogene point mutations distinguishes heritable  
 RT from nonheritable medullary thyroid carcinomas.";  
 RL Cancer 76:479-489(1995).  
 RN [20]  
 RP VARIANTS MEN2A SER-618; SER-620; ARG-634 AND TYR-634.  
 RX MEDLINE=95163936; PubMed=7860065;  
 RA Takiguchi-Shirahama S., Koyama K., Miyauchi A., Wakasugi T., Oishi S.,  
 RA Takami H., Hikiji K., Nakamura Y.;  
 RT "Germline mutations of the RET proto-oncogene in eight Japanese  
 RT patients with multiple endocrine neoplasia type 2A (MEN2A).";  
 RL Hum. Genet. 95:187-190(1995).  
 RN [21]  
 RP VARIANTS HSCR LEU-20; SER-93; GLN-330; TYR-609 AND ARG-620, AND  
 RP VARIANT CYS-982.  
 RX MEDLINE=95360000; PubMed=7633441;  
 RA Angrist M., Bolk S., Thiel B., Puffenberger E.G., Hofstra R.M.W.,  
 RA Buys C.H.C.M., Cass D.T., Chakravarti A.;  
 RT "Mutation analysis of the RET receptor tyrosine kinase in Hirschsprung  
 RT disease.";  
 RL Hum. Mol. Genet. 4:821-830(1995).  
 RN [22]  
 RP VARIANTS HSCR.  
 RX MEDLINE=96090258; PubMed=7581377;  
 RA Attie T., Pelet A., Ederly P., Eng C., Mulligan L.M., Amiel J.,  
 RA Boutrand L., Beldjord C., Nihoul-Fekete C., Munnich A., Ponder B.A.J.,  
 RA Lyonnet S.;  
 RT "Diversity of RET proto-oncogene mutations in familial and sporadic

RT Hirschsprung disease.";  
 RL Hum. Mol. Genet. 4:1381-1386(1995).  
 Query Match 24.3%; Score 304.5; DB 1; Length 1114;  
 Best Local Similarity 32.1%; Pred. No. 9.3e-21;  
 Matches 79; Conservative 44; Mismatches 78; Indels 45; Gaps 9;  
 QY 2 ELRQLSRVNHPIVILYGLACL--NPVCLVMEYAEAGSLYNVLHGAEPL-PYY----- 50  
 Db 775 EFNVLQVNHPIVILYGLACSDGGLLLIVEYAKYGLRGLRKRKVGFLGSGGSRN 834  
 QY 51 -----TAAHMSWCLQCSGVAYLHSMQPKALHRDLKPNLLVAGGTVLKI 98  
 Db 835 SSSLDHPDERALTMGLISFAWQISQMOYLAEMK--LVHRDLAARN-ILVAEGRMKI 890  
 QY 99 CDFTGACDI---QTHMTNKG--SAAWMAPEVEGSGNYSEKCDVFSWGIILWEIVTRRK 153  
 Db 891 SDFGLSRDVEEDSYKRSQGRIPVKWMAIESLFDHIYTTQSDVWSEFVLLWEIVT--- 946  
 QY 154 FDEIGG-----PAFRIMWAVHNGTRPPLIKNLPKPIESLMTCSKDPQSRPMEIVK 207  
 Db 947 ---LGNPYPGIPPERLENLKTGHRMERPDNCSEMYRLMLQCKQKQEPDKRPVFADISK 1003  
 QY 208 IMTHLM 213  
 Db 1004 DLEKWM 1009  
 RESULT 13  
 TEC MOUSE STANDARD; PRT; 630 AA.  
 ID AC P24604; Q9R1M9; Q9WVN0; Q9WVN1; Q9WVN2; Q9WVN3;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Tyrosine-protein kinase Tec (EC 2.7.1.112).  
 GN TEC.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=93149603; PubMed=7678927;  
 RA Mano H., Mano K., Tang B., Koshler M., Yi T., Gilbert D.J.,  
 RA Jenkins N.A., Copeland N.G., Ihle J.N.;  
 RT "Expression of a novel form of Tec kinase in hematopoietic cells and  
 RT mapping of the gene to chromosome 5 near Kit.";  
 RL Oncogene 8:417-424(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 3; 4; 5 AND 6).  
 RC STRAIN=129; PubMed=10343129;  
 RX MEDLINE=99276475; PubMed=10343129;  
 RA Merkel A.L., Atmosukarto I.I.C., Stevens K., Rathjen P.D.,  
 RA Booker G.W.;  
 RT "Splice variants of the mouse Tec gene are differentially expressed in  
 RT vivo.";  
 RL Cytogenet. Cell Genet. 84:132-139(1999).  
 RN [3]  
 RP SEQUENCE OF 82-630 FROM N.A. (ISOFORM 2).  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=91133729; PubMed=2284097;  
 RA Mano H., Ishikawa F., Nishida J., Hirai H., Takaku F.;  
 RT "A novel protein-tyrosine kinase, tec, is preferentially expressed in  
 RT liver.";  
 RL Oncogene 5:1781-1786(1990).  
 RN [4]  
 RP SEQUENCE OF 485-553 FROM N.A.  
 RX MEDLINE=90152381; PubMed=2482828;  
 RA Wilks A.F., Kurban R.R., Hovens C.M., Ralph S.J.;  
 RT "The application of the polymerase chain reaction to cloning members  
 RT of the protein tyrosine kinase family.";  
 RL Gene 85:67-74(1989).



[illegible]

[illegible]

DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DE 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Tyrosine-protein kinase ITK/TSK (EC 2.7.1.112) (T-cell-specific  
GN kinase) (II-2-inducible T-cell kinase) (Kinase EMT) (Kinase TLK).  
GN ITK OR TSK OR EMT OR TLK.  
OS Mus musculus (Mouse)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thymocytes;  
RX MEDLINE=93087493; PubMed=1280821;  
RA Siliciano J.D., Morrow T.A., Desiderio S.V.;  
RT "Itk, a T-cell-specific tyrosine kinase gene inducible by interleukin  
RL 2.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:11194-11198 (1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thymocytes;  
RX MEDLINE=93113848; PubMed=8421704;  
RA Heyeck S.D., Berg L.J.;  
RT "Developmental regulation of a murine T-cell-specific tyrosine kinase  
RL gene, Tsk.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:669-673 (1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CBA/J; TISSUE=Mast cells;  
RX MEDLINE=93236578; PubMed=8476425;  
RA Yamada N., Kawakami Y., Kimura H., Fukumachi H., Baier G.,  
RA Altman A., Kato T., Ingaki Y., Kawakami T.;  
RT "Structure and expression of novel protein-tyrosine kinases, Emb and  
RL Emt, in hematopoietic cells.";  
RL Biochem. Biophys. Res. Commun. 192:231-240 (1993).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Ogata M., Sawada M., Fujiwara H., Hamaoka T.;  
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP STRUCTURE BY NMR OF 160-236.  
RX MEDLINE=97138229; PubMed=8985255;  
RA Andreotti A.H., Bunnell S.C., Feng S., Berg L.J., Schreiber S.L.;  
RT "Regulatory intramolecular association in a tyrosine kinase of the  
RL Tec family.";  
RL Nature 385:93-97 (1997).  
RN [6]  
RP CHARACTERIZATION  
RX MEDLINE=95023908; PubMed=7524075;  
RA August A., Gibson S., Kawakami Y., Kawakami T., Mills G.B., Dupont B.;  
RT "CD28 is associated with and induces the immediate tyrosine  
RT phosphorylation and activation of the Tec family kinase ITK/EMT in  
RL the human Jurkat leukemic T-cell line.";  
RL Proc. Natl. Acad. Sci. U.S.A. 91:9347-9351 (1994).  
RN [7]  
RP CHARACTERIZATION  
RX MEDLINE=97098950; PubMed=8943565;  
RA King P.D., Sadra A., Han A., Liu X.-R., Sunder-Plassmann R.,  
RA Reinherz E.L., Dupont B.;  
RT "CD2 signaling in T cells involves tyrosine phosphorylation and  
RT activation of the Tec family kinase, EMT/ITK/TSK.";  
RL Int. Immunol. 8:1707-1714 (1996).  
RN [8]  
RP CHARACTERIZATION  
RX MEDLINE=20040393; PubMed=10570289;  
RA Ching K.A., Kawakami Y., Kawakami T., Tsoukas C.D.;  
RT "Emt/Itk associates with activated TCR complexes: role of the  
RT pleckstrin homology domain.";  
RL J. Immunol. 163:6006-6013 (1999).  
CC -!- FUNCTION: MAY PLAY A ROLE IN T-CELL DEVELOPMENT, POTENTIALLY IN  
CC THYMIC SELECTION.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBUNIT: LIGATION OF CD2, TCR OR CD28 INDUCES ACTIVATION AND

CC TYROSINE PHOSPHORYLATION OF ITK.  
CC -!- SUBCELLULAR LOCATION: LOCALIZES TO CELL SURFACE RECEPTORS IN THE  
CC PLASMA MEMBRANE AFTER STIMULATION WITH RESPECTIVE RECEPTORS (TCR,  
CC CD28, CD2) IN T-CELLS.  
CC -!- TISSUE SPECIFICITY: IS DETECTED IN THE THYMUS, LYMPH NODE AND VERY  
CC FAINTLY IN THE SPLEEN, BUT IS NOT DETECTED IN THE LIVER, LUNG,  
CC KIDNEY, HEART, BRAIN, INTESTINE OR TESTIS. EXPRESSED IN T-  
CC LYMPHOCYTES AND MAST CELLS. IT MAY ALSO BE EXPRESSED IN NATURAL  
CC KILLER CELLS.  
CC -!- DEVELOPMENTAL STAGE: IS PRESENT IN THE FETAL THYMUS AS EARLY AS  
CC DAY 14 OF GESTATION. THE LEVELS ARE 5- TO 10-FOLD HIGHER IN  
CC THYMOCYTES THAN IN PERIPHERAL T CELLS, AND INCREASE IN THE THYMUS  
CC DURING DEVELOPMENT FROM NEONATE TO ADULT.  
CC -!- INDUCTION: By interleukin-2.  
CC -!- DOMAIN: THE PH DOMAIN MEDIATES TARGETING OF ITK/TSK AND IS  
CC INDISPENSIBLE FOR THE ACTIVATION THROUGH TCR/CD3.  
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. TEC  
CC SUBFAMILY.  
CC -!- SIMILARITY: Contains 1 SH2 domain.  
CC -!- SIMILARITY: Contains 1 SH3 domain.  
CC -!- SIMILARITY: Contains 1 PH domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L00619; AAA39337.1; -;  
CC EMBL; L05631; AAA40518.1; -;  
CC EMBL; L10628; -; NOT ANNOTATED\_CDS.  
CC EMBL; D14042; BAA03129.1; -;  
CC PIR; A43030; A43030.  
CC PDB; 1AWJ; 14-JAN-98.  
CC PDB; 1LUK; 27-NOV-02.  
CC PDB; 1LUM; 27-NOV-02.  
CC PDB; 1LUN; 27-NOV-02.  
CC MGD; MGI:96621; Itk.  
CC InterPro; IPR001562; BTK.  
CC InterPro; IPR001849; PH.  
CC InterPro; IPR000719; Prot\_kinase.  
CC InterPro; IPR000980; SH2.  
CC InterPro; IPR001452; SH3.  
CC InterPro; IPR001245; Tyr\_pkinase.  
CC Pfam; PF00779; BTK; 1.  
CC Pfam; PF00169; PH; 1.  
CC Pfam; PF00069; Pkinase; 1.  
CC Pfam; PF00017; SH2; 1.  
CC Pfam; PF00018; SH3; 1.  
CC PRINTS; PR00401; SH2DOMAIN.  
CC PRINTS; PR00402; TECBTKDOMAIN.  
CC PRINTS; PR00109; TYRKINASE.  
CC ProDom; PD000001; Prot\_kinase; 1.  
CC ProDom; PD000093; SH2; 1.  
CC ProDom; PD000066; SH3; 1.  
CC SMART; SM00107; BTK; 1.  
CC SMART; SM00233; PH; 1.  
CC SMART; SM00252; SH2; 1.  
CC SMART; SM00326; SH3; 1.  
CC SMART; SM00219; Tyrc; 1.  
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
CC PROSITE; PS00109; PROTEIN KINASE TYR; 1.  
CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC PROSITE; PS50001; SH2; 1.  
CC PROSITE; PS50002; SH3; 1.  
CC PROSITE; PS50003; PH\_DOMAIN; 1.  
CC TRANSFERASE; Tyrosine-protein kinase; Phosphorylation;  
CC ATP-binding; SH2 domain; SH3 domain; 3D-structure.  
CC DOMAIN 4 117 PH.  
CC DOMAIN 177 237 SH3.  
CC DOMAIN 245 343 SH2.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 13:58:54 ; Search time 90.2825 Seconds  
(without alignments)  
1997.604 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAQRSLQEQQPSWTDD.....AEFYRLWSVDHGQSVVTAP 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1589859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1589859

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2580	100.0	504	13	US-10-123-427-2
2	2580	100.0	504	13	US-10-156-895-2
3	2580	100.0	504	14	US-10-384-743-2
4	2580	100.0	504	17	US-10-820-583A-10
5	2580	100.0	517	13	US-10-156-895-11
6	2580	100.0	517	14	US-10-384-743-11
7	2577	99.9	513	13	US-10-156-895-43
8	2577	99.9	513	14	US-10-384-743-43
9	2575	99.8	504	13	US-10-123-427-6
10	409	15.9	84	9	US-09-925-300-1270
11	391	15.2	180	17	US-10-425-115-264407
12	365	14.1	70	9	US-09-864-761-34065
13	261	10.1	51	9	US-09-864-761-34067

14	249.5	9.7	252	15	US-10-072-012-854	Sequence 854, App
15	244.5	9.5	260	15	US-10-072-012-853	Sequence 853, App
16	218	8.4	338	9	US-09-801-267-5	Sequence 5, Appli
17	218	8.4	338	14	US-10-170-789-36	Sequence 36, Appl
18	208	8.1	338	15	US-10-424-599-226401	Sequence 226401,
19	208	8.1	338	15	US-10-424-599-226412	Sequence 226412,
20	206	8.0	432	15	US-10-425-114-59312	Sequence 59312, A
21	205	7.9	274	9	US-09-860-351-4	Sequence 4, Appli
22	205	7.9	300	9	US-09-801-367-4	Sequence 4, Appli
23	205	7.9	300	14	US-10-170-789-35	Sequence 35, Appl
24	203.5	7.9	284	16	US-10-437-963-151472	Sequence 151472,
25	202	7.8	370	17	US-10-425-115-314480	Sequence 314480,
26	202	7.8	416	15	US-10-425-114-72766	Sequence 72766, A
27	201	7.8	312	15	US-10-424-599-219880	Sequence 219880,
28	200.5	7.8	354	17	US-10-425-115-189784	Sequence 189784,
29	198	7.7	290	17	US-10-425-115-288785	Sequence 288785,
30	198	7.7	311	15	US-10-425-114-42276	Sequence 42276, A
31	198	7.7	311	15	US-10-425-114-55347	Sequence 55347, A
32	198	7.7	311	15	US-10-425-114-67199	Sequence 67199, A
33	197.5	7.7	327	16	US-10-437-963-125417	Sequence 125417,
34	197.5	7.7	356	15	US-10-425-114-39847	Sequence 39847, A
35	197.5	7.7	423	15	US-10-425-114-49889	Sequence 49889, A
36	196.5	7.6	356	15	US-10-424-599-208454	Sequence 208454,
37	194.5	7.5	363	15	US-10-424-599-251587	Sequence 251587,
38	194.5	7.5	384	15	US-10-425-114-55810	Sequence 55810, A
39	194	7.5	290	16	US-10-437-963-167104	Sequence 167104,
40	193	7.5	362	16	US-10-437-963-144908	Sequence 144908,
41	192	7.4	285	15	US-10-424-599-206342	Sequence 206342,
42	191.5	7.4	284	17	US-10-425-115-298390	Sequence 298390,
43	191.5	7.4	314	15	US-10-425-115-219869	Sequence 219869,
44	191.5	7.4	321	15	US-10-425-114-58854	Sequence 58854, A
45	191	7.4	284	16	US-10-767-701-45122	Sequence 45122, A

## ALIGNMENTS

RESULT 1  
US-10-123-427-2  
; Sequence 2, Application US/10123427  
; Publication No. US20020119525A1  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiko  
; INVENTOR: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/123,427  
; FILING DATE: 17-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIORITY INFORMATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/406,854  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 504 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-123-427-2

Query Match 100.0%; Score 2580; DB 13; Length 504;  
Best Local Similarity 100.0%; Pred. No. 4.5e-200;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAAQRSLQSQPQSWTDDPLCHLSGVGSASNSYSADGKGTESHPPEDSWLKFRSEN 60  
DB 1 MAAQRSLQSQPQSWTDDPLCHLSGVGSASNSYSADGKGTESHPPEDSWLKFRSEN 60  
QY 61 NCFLYGVFNGYDGNRVTFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSELES 120  
DB 61 NCFLYGVFNGYDGNRVTFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSELES 120  
QY 121 IDALLAEKASLQSQPQSWTDDPLCHLSGVGSASNSYSADGKGTESHPPEDSWLKFRSEN 180  
DB 121 IDALLAEKASLQSQPQSWTDDPLCHLSGVGSASNSYSADGKGTESHPPEDSWLKFRSEN 180  
QY 181 ANVGTRNALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240  
DB 181 ANVGTRNALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240  
QY 241 RIGDYKVYGYTDDILLSAKSKPIIAEPIHGAQPLDGVTFGLVMSSEGLYKALEAAH 300  
DB 241 RIGDYKVYGYTDDILLSAKSKPIIAEPIHGAQPLDGVTFGLVMSSEGLYKALEAAH 300  
QY 301 GPGQANQEIATMDTEFAKQTSIDAVAQAVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360  
DB 301 GPGQANQEIATMDTEFAKQTSIDAVAQAVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360  
QY 361 LVNFGYPLGEMSOPTSPAPAGGRVYVPSVYSSAQSTKSTVTLVMPSSQGMVNG 420  
DB 361 LVNFGYPLGEMSOPTSPAPAGGRVYVPSVYSSAQSTKSTVTLVMPSSQGMVNG 420  
QY 421 AHSASTLDEATPTLNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
DB 421 AHSASTLDEATPTLNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
QY 481 YVDFAEFYRLWSVDHGEQSVVTAP 504  
DB 481 YVDFAEFYRLWSVDHGEQSVVTAP 504

RESULT 2  
US-10-158-895-2  
Sequence 2, Application US/10158895  
Publication No. US20020155624A1  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/158,895  
PRIOR FILING DATE: 2002-06-03  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188

PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 504  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-158-895-2  
Query Match 100.0%; Score 2580; DB 13; Length 504;  
Best Local Similarity 100.0%; Pred. No. 4.5e-200;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAAQRSLQSQPQSWTDDPLCHLSGVGSASNSYSADGKGTESHPPEDSWLKFRSEN 60  
DB 1 MAAQRSLQSQPQSWTDDPLCHLSGVGSASNSYSADGKGTESHPPEDSWLKFRSEN 60  
QY 61 NCFLYGVFNGYDGNRVTFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSELES 120  
DB 61 NCFLYGVFNGYDGNRVTFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSELES 120  
QY 121 IDALLAEKASLQSQPQSWTDDPLCHLSGVGSASNSYSADGKGTESHPPEDSWLKFRSEN 180  
DB 121 IDALLAEKASLQSQPQSWTDDPLCHLSGVGSASNSYSADGKGTESHPPEDSWLKFRSEN 180  
QY 181 ANVGTRNALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240  
DB 181 ANVGTRNALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240  
QY 241 RIGDYKVYGYTDDILLSAKSKPIIAEPIHGAQPLDGVTFGLVMSSEGLYKALEAAH 300  
DB 241 RIGDYKVYGYTDDILLSAKSKPIIAEPIHGAQPLDGVTFGLVMSSEGLYKALEAAH 300  
QY 301 GPGQANQEIATMDTEFAKQTSIDAVAQAVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360  
DB 301 GPGQANQEIATMDTEFAKQTSIDAVAQAVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360  
QY 361 LVNFGYPLGEMSOPTSPAPAGGRVYVPSVYSSAQSTKSTVTLVMPSSQGMVNG 420  
DB 361 LVNFGYPLGEMSOPTSPAPAGGRVYVPSVYSSAQSTKSTVTLVMPSSQGMVNG 420  
QY 421 AHSASTLDEATPTLNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
DB 421 AHSASTLDEATPTLNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
QY 481 YVDFAEFYRLWSVDHGEQSVVTAP 504  
DB 481 YVDFAEFYRLWSVDHGEQSVVTAP 504

RESULT 3  
US-10-384-743-2  
Sequence 2, Application US/10384743  
Publication No. US2003016228A1  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/10/384,743  
PRIOR FILING DATE: 2003-03-11  
CURRENT FILING DATE: 2003-03-11  
PRIOR APPLICATION NUMBER: US/09/529,279  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 504  
TYPE: PRT

```

; ORGANISM: Homo sapiens
US-10-384-743-2

Query Match      100.0%; Score 2580; DB 14; Length 504;
Best Local Similarity 100.0%; Pred. No. 4.5e-200;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60

QY 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDIVERSFLES 120
Db 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDIVERSFLES 120

QY 121 IDDALAEKASLOSQLPQGVPOHQIPPOYQKILERLKTILERISGGAMAVAVLLNNKLYV 180
Db 121 IDDALAEKASLOSQLPQGVPOHQIPPOYQKILERLKTILERISGGAMAVAVLLNNKLYV 180

QY 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICGQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICGQEST 240

QY 241 RRGIDYKVKYGYTDIDLLSAAKSPKPIIAEPEIHAQAQPLDGVTFGLVLMSEGLYKALEAAH 300
Db 241 RRGIDYKVKYGYTDIDLLSAAKSPKPIIAEPEIHAQAQPLDGVTFGLVLMSEGLYKALEAAH 300

QY 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360

QY 361 LVNFGYPLGEMSOPTSPAPAGGRVYPVSVYSSAQSTSKTSVTLSLWMPGQGMVNG 420
Db 361 LVNFGYPLGEMSOPTSPAPAGGRVYPVSVYSSAQSTSKTSVTLSLWMPGQGMVNG 420

QY 421 AHSASTLDEATPTLTNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480

QY 481 YVDFAEFYRLWSVDHGEOQSVVTAP 504
Db 481 YVDFAEFYRLWSVDHGEOQSVVTAP 504

RESULT 4
US-10-820-583A-10
; Sequence 10, Application US/10820583A
; Publication No. US20040242461A1
; GENERAL INFORMATION:
; APPLICANT: Schneider, Michael
; APPLICANT: Oh, Hidemasa
; TITLE OF INVENTION: Modulators of Telomere Stability
; CURRENT APPLICATION NUMBER: US/10/820,583A
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: US 60/461,095
; PRIOR FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 504
; TYPE: PRT
; ORGANISM: HUMAN
US-10-820-583A-10

Query Match      100.0%; Score 2580; DB 17; Length 504;
Best Local Similarity 100.0%; Pred. No. 4.5e-200;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60

QY 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDIVERSFLES 120
Db 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDIVERSFLES 120

QY 121 IDDALAEKASLOSQLPQGVPOHQIPPOYQKILERLKTILERISGGAMAVAVLLNNKLYV 180
Db 121 IDDALAEKASLOSQLPQGVPOHQIPPOYQKILERLKTILERISGGAMAVAVLLNNKLYV 180

QY 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICGQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICGQEST 240

QY 241 RRGIDYKVKYGYTDIDLLSAAKSPKPIIAEPEIHAQAQPLDGVTFGLVLMSEGLYKALEAAH 300
Db 241 RRGIDYKVKYGYTDIDLLSAAKSPKPIIAEPEIHAQAQPLDGVTFGLVLMSEGLYKALEAAH 300

QY 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360

QY 361 LVNFGYPLGEMSOPTSPAPAGGRVYPVSVYSSAQSTSKTSVTLSLWMPGQGMVNG 420
Db 361 LVNFGYPLGEMSOPTSPAPAGGRVYPVSVYSSAQSTSKTSVTLSLWMPGQGMVNG 420

QY 421 AHSASTLDEATPTLTNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480

QY 481 YVDFAEFYRLWSVDHGEOQSVVTAP 504
Db 481 YVDFAEFYRLWSVDHGEOQSVVTAP 504

```

```

QY 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDIVERSFLES 120
Db 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDIVERSFLES 120

QY 121 IDDALAEKASLOSQLPQGVPOHQIPPOYQKILERLKTILERISGGAMAVAVLLNNKLYV 180
Db 121 IDDALAEKASLOSQLPQGVPOHQIPPOYQKILERLKTILERISGGAMAVAVLLNNKLYV 180

QY 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICGQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICGQEST 240

QY 241 RRGIDYKVKYGYTDIDLLSAAKSPKPIIAEPEIHAQAQPLDGVTFGLVLMSEGLYKALEAAH 300
Db 241 RRGIDYKVKYGYTDIDLLSAAKSPKPIIAEPEIHAQAQPLDGVTFGLVLMSEGLYKALEAAH 300

QY 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360

QY 361 LVNFGYPLGEMSOPTSPAPAGGRVYPVSVYSSAQSTSKTSVTLSLWMPGQGMVNG 420
Db 361 LVNFGYPLGEMSOPTSPAPAGGRVYPVSVYSSAQSTSKTSVTLSLWMPGQGMVNG 420

QY 421 AHSASTLDEATPTLTNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNQSPPTLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480

QY 481 YVDFAEFYRLWSVDHGEOQSVVTAP 504
Db 481 YVDFAEFYRLWSVDHGEOQSVVTAP 504

RESULT 5
US-10-158-895-11
; Sequence 11, Application US/10158895
; Publication No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-11

Query Match      100.0%; Score 2580; DB 13; Length 517;
Best Local Similarity 100.0%; Pred. No. 4.6e-200;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60

QY 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDIVERSFLES 120
Db 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDIVERSFLES 120

QY 121 IDDALAEKASLOSQLPQGVPOHQIPPOYQKILERLKTILERISGGAMAVAVLLNNKLYV 180
Db 121 IDDALAEKASLOSQLPQGVPOHQIPPOYQKILERLKTILERISGGAMAVAVLLNNKLYV 180

```

Db	121	IDDALAEKASLSQSLPEGVPQHLPPOYQKILERIKTLEREISGGAMAVAVLNNKLYV	180
Qy	181	ANVTNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST	240
Db	181	ANVTNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST	240
Qy	241	RRIGDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH	300
Db	241	RRIGDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH	300
Qy	301	GPQANQIEAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL	360
Db	301	GPQANQIEAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL	360
Qy	361	LVRNFGYPLGEMSQTPSPAPAAGGRVVPVSPYSSAOSTSKTSVTLSLWMP SQQMVG	420
Db	361	LVRNFGYPLGEMSQTPSPAPAAGGRVVPVSPYSSAOSTSKTSVTLSLWMP SQQMVG	420
Qy	421	AHSASTLDEATPTLNQSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP	480
Db	421	AHSASTLDEATPTLNQSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP	480
Qy	481	YVDFAEFYRLMSVDHGEQSVVTPAP	504
Db	481	YVDFAEFYRLMSVDHGEQSVVTPAP	504
RESULT 6			
US-10-384-743-11			
; Sequence 11, Application US/10384743			
; Publication No. US20030162228A1			
; GENERAL INFORMATION:			
; APPLICANT: ONO, KOICHIRO			
; APPLICANT: TSUCHIYA, MASAYUKI			
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES			
; FILE REFERENCE: 053466/0278			
; CURRENT APPLICATION NUMBER: US/10/384,743			
; PRIOR FILING DATE: 2003-03-11			
; PRIOR APPLICATION NUMBER: US/09/529,279			
; PRIOR FILING DATE: 2000-04-11			
; PRIOR APPLICATION NUMBER: PCT/JP98/04796			
; PRIOR FILING DATE: 1998-10-22			
; PRIOR APPLICATION NUMBER: JP 9/290188			
; NUMBER OF SEQ ID NOS: 48			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 11			
; LENGTH: 517			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-384-743-11			
Query Match 100.0%; Score 2580; DB 14; Length 517;			
Best Local Similarity 100.0%; Pred. No. 4.6e-200;			
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAAQRSLQSEQQPSWTDPLCHLSGVGSANRSYSADGKGTESHPPEDSMLKFRSEN	60
Db	1	MAAQRSLQSEQQPSWTDPLCHLSGVGSANRSYSADGKGTESHPPEDSMLKFRSEN	60
Qy	61	NCFLYGVFNGYDGNRVNFFVAQRLSAELLIGQLNAEHAEDVRVLLQAEDVVERSFLES	120
Db	61	NCFLYGVFNGYDGNRVNFFVAQRLSAELLIGQLNAEHAEDVRVLLQAEDVVERSFLES	120
Qy	121	IDDALAEKASLSQSLPEGVPQHLPPOYQKILERIKTLEREISGGAMAVAVLNNKLYV	180
Db	121	IDDALAEKASLSQSLPEGVPQHLPPOYQKILERIKTLEREISGGAMAVAVLNNKLYV	180
Qy	181	ANVTNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST	240
Db	181	ANVTNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST	240
Qy	241	RRIGDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH	300
Db	241	RRIGDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH	300
Qy	301	GPQANQIEAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL	360
Db	301	GPQANQIEAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL	360
US-10-158-895-43			
; Sequence 43, Application US/10158895			
; Publication No. US20020155624A1			
; GENERAL INFORMATION:			
; APPLICANT: ONO, KOICHIRO			
; APPLICANT: OHTOMO, TOSHIHIKO			
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES			
; FILE REFERENCE: 053466/0278			
; CURRENT APPLICATION NUMBER: US/10/158,895			
; PRIOR FILING DATE: 2002-06-03			
; PRIOR APPLICATION NUMBER: US/09/529,279			
; PRIOR FILING DATE: 2000-04-11			
; PRIOR APPLICATION NUMBER: PCT/JP98/04796			
; PRIOR FILING DATE: 1998-10-22			
; PRIOR APPLICATION NUMBER: JP 9/290188			
; PRIOR FILING DATE: 1997-10-22			
; NUMBER OF SEQ ID NOS: 48			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 43			
; LENGTH: 513			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-158-895-43			
Query Match 99.9%; Score 2577; DB 13; Length 513;			
Best Local Similarity 99.8%; Pred. No. 8e-200;			
Matches 503; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAAQRSLQSEQQPSWTDPLCHLSGVGSANRSYSADGKGTESHPPEDSMLKFRSEN	60
Db	10	MAAQRSLQSEQQPSWTDPLCHLSGVGSANRSYSADGKGTESHPPEDSMLKFRSEN	69
Qy	61	NCFLYGVFNGYDGNRVNFFVAQRLSAELLIGQLNAEHAEDVRVLLQAEDVVERSFLES	120
Db	70	NCFLYGVFNGYDGNRVNFFVAQRLSAELLIGQLNAEHAEDVRVLLQAEDVVERSFLES	129
Qy	121	IDDALAEKASLSQSLPEGVPQHLPPOYQKILERIKTLEREISGGAMAVAVLNNKLYV	180
Db	130	IDDALAEKASLSQSLPEGVPQHLPPOYQKILERIKTLEREISGGAMAVAVLNNKLYV	189
Qy	181	ANVTNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST	240
Db	190	ANVTNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST	249
Qy	241	RRIGDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH	300
Db	250	RRIGDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH	309
Qy	301	GPQANQIEAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL	360



Db 310 GPCQANQEIAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTEFASGGERARFCPRHEDMTL 369  
QY 361 LVNFGYPLGMSQPTSPAPAGGRVYVPVYSSAQSTSKTSVTLVMPQSQGMVNG 420  
Db 370 LVNFGYPLGMSQPTSPAPAGGRVYVPVYSSAQSTSKTSVTLVMPQSQGMVNG 429  
QY 421 AHSASTLDEATPILTNQSPILTLQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVPE 480  
Db 430 AHSASTLDEATPILTNQSPILTLQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVPE 489  
QY 481 YVDFAEFYRLWSVDHGEQSVVVTAP 504  
Db 490 YVDFAEFYRLWSVDHGEQSVVVTAP 513  
RESULT 8  
US-10-384-743-43  
; Sequence 43, Application US/10384743  
; Publication No. US20030162228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; PRIOR FILING DATE: 2003-03-11  
; PRIOR FILING DATE: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 43  
; LENGTH: 513  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-384-743-43

Query Match 99.9%; Score 2577; DB 14; Length 513;  
Best Local Similarity 99.8%; Pred. No. 8e-200;  
Matches 503; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAQRSLQSQEQSPSWTDDPLCHLSGVGSASNRYSADGKGTEGHPEDSWLKRSEN 60  
Db 10 MAQRSLQSQEQSPSWTDDPLCHLSGVGSASNRYSADGKGTEGHPEDSWLKRSEN 69  
QY 61 NCFLYGVFNQYDGNRVNFEVQRLSAEALLGQNAEHAEDVRRVLLQAFDIVERSFLES 120  
Db 70 NCFLYGVFNQYDGNRVNFEVQRLSAEALLGQNAEHAEDVRRVLLQAFDIVERSFLES 129  
QY 121 IDDAEAKASQSQPEGVQHQLPQYQKILRLKTLEREISGGAMAVVAVLNKLYV 180  
Db 130 IDDAEAKASQSQPEGVQHQLPQYQKILRLKTLEREISGGAMAVVAVLNKLYV 189  
QY 181 ANVGNRALLCKSTVDGLQVTLQNVDTHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240  
Db 190 ANVGNRALLCKSTVDGLQVTLQNVDTHTTENEDELFRSLQGLDAGKIKQVGIICQEST 249  
QY 241 RRGIVKVKYGYTDDLLSAASKPIIARPEIHGAQPLDGVTFGLVLMSEGLYKALEAH 300  
Db 250 RRGIVKVKYGYTDDLLSAASKPIIARPEIHGAQPLDGVTFGLVLMSEGLYKALEAH 309  
QY 301 GPCQANQEIAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTEFASGGERARFCPRHEDMTL 360  
Db 310 GPCQANQEIAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTEFASGGERARFCPRHEDMTL 369  
QY 361 LVNFGYPLGMSQPTSPAPAGGRVYVPVYSSAQSTSKTSVTLVMPQSQGMVNG 420  
Db 370 LVNFGYPLGMSQPTSPAPAGGRVYVPVYSSAQSTSKTSVTLVMPQSQGMVNG 429

QY 421 AHSASTLDEATPILTNQSPILTLQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVPE 480  
Db 430 AHSASTLDEATPILTNQSPILTLQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVPE 489  
QY 481 YVDFAEFYRLWSVDHGEQSVVVTAP 504  
Db 490 YVDFAEFYRLWSVDHGEQSVVVTAP 513  
RESULT 9  
US-10-123-427-6  
; Sequence 6, Application US/10123427  
; Publication No. US20020119525A1  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/123,427  
; FILING DATE: 17-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/406,854  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-10-123-427-6

Query Match 99.8%; Score 2575; DB 13; Length 504;  
Best Local Similarity 99.8%; Pred. No. 1.1e-199;  
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MAQRSLQSQEQSPSWTDDPLCHLSGVGSASNRYSADGKGTEGHPEDSWLKRSEN 60  
Db 1 MAQRSLQSQEQSPSWTDDPLCHLSGVGSASNRYSADGKGTEGHPEDSWLKRSEN 60  
QY 61 NCFLYGVFNQYDGNRVNFEVQRLSAEALLGQNAEHAEDVRRVLLQAFDIVERSFLES 120  
Db 61 NCFLYGVFNQYDGNRVNFEVQRLSAEALLGQNAEHAEDVRRVLLQAFDIVERSFLES 120  
QY 121 IDDAEAKASQSQPEGVQHQLPQYQKILRLKTLEREISGGAMAVVAVLNKLYV 180

Db 121 IDDAAEKASLQSPGVPQHQLPPOYQKILERLKLTEREISGAMAVAVLLNNKLIY 180  
Qy 181 ANVTNNALLCKSVTDGLQVTLNVHDHTTENEDELFRLSQLDAGKIKQVGIICGQEST 240  
Db 181 ANVTNNALLCKSVTDGLQVTLNVHDHTTENEDELFRLSQLDAGKIKQVGIICGQEST 240  
Qy 241 RRIIGDYKVKYGTIDILLSAKSPKIIAEPHGAQPLDGVTFVLVMSGLYKALEAAH 300  
Db 241 RRIIGDYKVKYGTIDILLSAKSPKIIAEPHGAQPLDGVTFVLVMSGLYKALEAAH 300  
Qy 301 GPGQANQEIAMIDTEFAKQTSLDAAVAQAVVDRVKRIHSDTFASGGERARCPRHEDMTL 360  
Db 301 GPGQANQEIAMIDTEFAKQTSLDAAVAQAVVDRVKRIHSDTFASGGERARCPRHEDMTL 360  
Qy 361 LVRNFGYPLGEMSOPTSPAPAAAGRVVPVSPYSSAQSTKSTVTLVMPSCQMVNG 420  
Db 361 LVRNFGYPLGEMSOPTSPAPAAAGRVVPVSPYSSAQSTKSTVTLVMPSCQMVNG 420  
Qy 421 AHSASTLDEATPTLTNQSPPTLTQSTNTHQTSSSSSDGGLFRSPAHSLPPGEDGRVEP 480  
Db 421 AHSASTLDEATPTLTNQSPPTLTQSTNTHQTSSSSSDGGLFRSPAHSLPPGEDGRVEP 480  
Qy 481 YVDFAEFYRLKSVDHGQSVVTAP 504  
Db 481 YVDFAEFYRLKSVDHGQSVVTAP 504

RESULT 10  
US-09-925-300-1270  
; Sequence 1270, Application US/09925300  
; Patent No. US20020151681A1  
; GENERAL INFORMATION:  
; APPLICANT: Craig Rosen,  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA101  
; CURRENT APPLICATION NUMBER: US/09/925,300  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05988  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1890  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 1270  
; LENGTH: 84  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; LOCATION: (38)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-300-1270

Query Match 15.9%; Score 409; DB 9; Length 84;  
Best Local Similarity 97.5%; Pred. No. 2.3e-25;  
Matches 78; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 425 STLDEATPTLTNQSPPTLTQSTNTHQTSSSSSDGGLFRSPAHSLPPGEDGRVEPYVDF 484  
Db 5 ATLDEATPTLTNQSPPTLTQSTNTHQTSSSSSDGGLFRSPAHSLPPGEDGRVEPYVDF 64  
Qy 485 AEFYRLKSVDHGQSVVTAP 504  
Db 65 AEFYRLKSVDHGQSVVTAP 84

RESULT 11  
US-10-425-115-264407  
; Sequence 264407, Application US/10425115  
; Publication No. US2004021472A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 264407  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MFT4577\_172752C.1.pep  
US-10-425-115-264407  
Query Match 15.2%; Score 391; DB 17; Length 180;  
Best Local Similarity 57.2%; Pred. No. 2.1e-23;  
Matches 87; Conservative 10; Mismatches 31; Indels 24; Gaps 3;  
Qy 258 LSAKSPKIIAEPHGAQPLDGVTFVLVMSGLYKALEAAHSGQANQEIAMIDTEF 317  
Db 25 LSKTSLAPIIV-----YVKISSPKVQLRIKSRGKSOAKHLNVQWVAADK 69  
Qy 318 AKQ-----TSLDAVAQAVVDRVKRIHSDTFASGGERARCPRHEDMTLVRNFGYPLG 370  
Db 70 LAQCPELDFVILDAVAQAVVDRVKRIHSDTFASGGERARCPRHEDMTLLVRNFGYPLG 129  
Qy 371 EMSOPTSPAPAAAGRVVPVSPYSSAQSTK 402  
Db 130 EMSQTPSPAP--GGRVVPVSPYSSAQSTR 159  
RESULT 12  
US-09-864-761-34065  
; Sequence 34065, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34065
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z83845.14
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EST HUMAN HIT: BE746542.1, EVALUE 5.00e-35
; OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUE 4.00e-36
US-09-864-761-34065

Query Match 14.1%; Score 365; DB 9; Length 70;
Best Local Similarity 100.0%; Pred. No. 6.4e-22;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 312 MIDTEFAKTSIDAVAQAQVVDVKRTHSTFASGGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKTSIDAVAQAQVVDVKRTHSTFASGGGERARFCPRHEDMTLLVRNFGYPLGE 60

QY 372 MSQPTSPAP 381
Db 61 MSQPTSPAP 70

RESULT 13
US-09-864-761-34067
; Sequence 34067, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34067
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z83845.14
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
; OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUE 7.00e-24
; OTHER INFORMATION: EST_HUMAN HIT: AL118967.1, EVALUE 9.00e-23
US-09-864-761-34067

Query Match 10.1%; Score 261; DB 9; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e-13;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 SENNCFLXGVFNGYDGNRVTFVQAQLSAELLGQLNAPHAADVRRLVLLQ 108
Db 1 SENNCFLXGVFNGYDGNRVTFVQAQLSAELLGQLNAPHAADVRRLVLLQ 51

RESULT 14
US-10-072-012-854
; Sequence 854, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zethusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
```

Wed Dec 22 14:29:07 2004

; Sequence 853, Application US/10072012	
; Publication No. US20040033493A1	
; GENERAL INFORMATION:	
; APPLICANT: Tchernev, Velizar	
; APPLICANT: Spytek, Kimberly	
; APPLICANT: Zerhusen, Bryan	
; APPLICANT: Patturajan, Meera	
; APPLICANT: Shimkets, Richard	
; APPLICANT: Li, Li	
; APPLICANT: Gangolli, Esha	
; APPLICANT: Padigaru, Muralidhara	
; APPLICANT: Anderson, David W.	
; APPLICANT: Rastelli, Luca	
; APPLICANT: Miller, Charles B.	
; APPLICANT: Gerlach, Valerie	
; APPLICANT: Taupier Jr, Raymond J.	
; APPLICANT: Gusev, Vladimir Y.	
; APPLICANT: Colman, Steven D.	
; APPLICANT: Wolenc, Adam R.	
; APPLICANT: Pena, Carol E. A.	
; APPLICANT: Furtak, Katarzyna	
; APPLICANT: Grose, William M.	
; APPLICANT: Alsobrook II, John P.	
; APPLICANT: Lepley, Denise M.	
; APPLICANT: Rieger, Daniel K.	
; APPLICANT: Burgess, Catherine E.	
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same	
; FILE REFERENCE: 21402-258	
; CURRENT APPLICATION NUMBER: US/10/072,012	
; CURRENT FILING DATE: 2002-01-31	
; PRIOR APPLICATION NUMBER: 60/265,102	
; PRIOR FILING DATE: 2001-01-30	
; PRIOR APPLICATION NUMBER: 60/265,514	
; PRIOR FILING DATE: 2001-01-31	
; PRIOR APPLICATION NUMBER: 60/265,517	
; PRIOR FILING DATE: 2001-01-31	
; PRIOR APPLICATION NUMBER: 60/265,412	
; PRIOR FILING DATE: 2001-01-31	
; PRIOR APPLICATION NUMBER: 60/265,395	
; PRIOR FILING DATE: 2001-01-31	
; PRIOR APPLICATION NUMBER: 60/266,406	
; PRIOR FILING DATE: 2001-02-02	
; PRIOR APPLICATION NUMBER: 60/266,767	
; PRIOR FILING DATE: 2001-02-05	
; PRIOR APPLICATION NUMBER: 60/267,057	
; PRIOR FILING DATE: 2001-02-07	
; PRIOR APPLICATION NUMBER: 60/266,975	
; PRIOR FILING DATE: 2001-02-07	
; PRIOR APPLICATION NUMBER: 60/267,459	
; PRIOR FILING DATE: 2001-02-08	
; Remaining Prior Application data removed - See File Wrapper or PALM.	
; NUMBER OF SEQ ID NOS: 1391	
; SOFTWARE: Patent in Ver. 2.1	
; SEQ ID NO 853	
; LENGTH: 260	
; TYPE: PRT	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence:	
; OTHER INFORMATION: Serine/threonine phosphatases, family 2C,	
; OTHER INFORMATION: catalytic domain	
US-10-072-012-853	
Query Match 9.5%; Score 244.5; DB 15; Length 260;	
Best Local Similarity 27.9%; Pred. No. 2.6e-11;	
Matches 81; Conservative 59; Mismatches 77; Indels 73; Gaps 14;	
QY 48 PPDSWL---KFRSENNCFLYGVNGYDGNRVTFVAQRLSAEILLGQLNAEHAADVRR 104	
Db 23 PMEDAHVITPDLSDGSDSGGFGFVFDGHHGSGAAKFKSKNL-PRILABEL----- 70	
QY 105 VLLQAFDVVERSFLESIDDAEAKASLQSLPEGVPQHLPPQYOKILERLKTLEIFSG 164	
RESULT 15	
US-10-072-012-853	
; APPLICANT: Gusev, Vladimir Y.	
; APPLICANT: Colman, Steven D.	
; APPLICANT: Wolenc, Adam R.	
; APPLICANT: Pena, Carol E. A.	
; APPLICANT: Furtak, Katarzyna	
; APPLICANT: Grose, William M.	
; APPLICANT: Alsobrook II, John P.	
; APPLICANT: Lepley, Denise M.	
; APPLICANT: Rieger, Daniel K.	
; APPLICANT: Burgess, Catherine E.	
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same	
; FILE REFERENCE: 21402-258	
; CURRENT APPLICATION NUMBER: US/10/072,012	
; CURRENT FILING DATE: 2002-01-31	
; PRIOR APPLICATION NUMBER: 60/265,102	
; PRIOR FILING DATE: 2001-01-30	
; PRIOR APPLICATION NUMBER: 60/265,514	
; PRIOR FILING DATE: 2001-01-31	
; PRIOR APPLICATION NUMBER: 60/265,517	
; PRIOR FILING DATE: 2001-01-31	
; PRIOR APPLICATION NUMBER: 60/265,412	
; PRIOR FILING DATE: 2001-01-31	
; PRIOR APPLICATION NUMBER: 60/265,395	
; PRIOR FILING DATE: 2001-01-31	
; PRIOR APPLICATION NUMBER: 60/266,406	
; PRIOR FILING DATE: 2001-02-02	
; PRIOR APPLICATION NUMBER: 60/266,767	
; PRIOR FILING DATE: 2001-02-05	
; PRIOR APPLICATION NUMBER: 60/267,057	
; PRIOR FILING DATE: 2001-02-07	
; PRIOR APPLICATION NUMBER: 60/266,975	
; PRIOR FILING DATE: 2001-02-07	
; PRIOR APPLICATION NUMBER: 60/267,459	
; PRIOR FILING DATE: 2001-02-08	
; Remaining Prior Application data removed - See File Wrapper or PALM.	
; NUMBER OF SEQ ID NOS: 1391	
; SOFTWARE: Patent in Ver. 2.1	
; SEQ ID NO 854	
; LENGTH: 252	
; TYPE: PRT	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence: Protein	
; OTHER INFORMATION: phosphatase 2C Consensus Sequence	
US-10-072-012-854	
Query Match 9.7%; Score 249.5; DB 15; Length 252;	
Best Local Similarity 26.0%; Pred. No. 9.9e-12;	
Matches 76; Conservative 49; Mismatches 80; Indels 87; Gaps 10;	
QY 30 GSASNRYSADGKGTSHPEDSWLKFRSENNCFLYGVNGYDGNRVTFVAQRLSAELL 89	
Db 22 GKNLSSGGKDSK-----FFAVFDGHHGSGAAKFKSKNL-PRILABEL----- 61	
QY 90 LQQLNAEHAADVRRVLLQAFDVVERSFLESIDDAEAKASLQSLPEGVPQHLPPQY 149	
Db 62 LRKSFPELDLE-----NALKSFLESTDEELRSSAA----- 93	
QY 150 KILERLKTLEIFSGGAMAVVALLNNKLVANVTNRALLCKSTVDGLQVLTQINVDHTT 209	
Db 94 -----NTDLSGSGTAVVALIRGNKLVANVGDSRAVLCR---NGNAIKQLTDEHKP 141	
QY 210 ENEDELFRSLQGLDAGIKQVGHICQES--TRRIGYKVKYGYTDIDLLSAKSKPIIA 268	
Db 142 SNEDERRRIEGAG---GFVSRNGRVNGVLAVSRAFGDFELKPGVL-----QPVTA 188	
QY 269 EPEI--HGAQPLDGVTCFLVMSGLYKALEAAHGPQAGQANCEIAAMIDTEFA 318	
Db 189 EPDVTSHKITPSD---BFLILASGLDVL-----SDQEVVDIVRSLS 229	
RESULT 15	
US-10-072-012-853	

Db 71 -----IKDKDEDEDVEDAL-RKAFRLTD-----EEILBELESLEDQRS- 107  
QY 165 GAMAVVAVLLNNKLYVANVGTNRALLCKSTVDGLQVLTQNLVVDHTTENEDLRLSQLG-- 222  
Db 108 GTTAVVALIRGNKUYVANVGDSRAVLCRNG---KAVQLTEDHKPSNEDERERIREAGGF 163  
QY 223 LDAGKIKQVGIICQESTRRIGDYKVYGYTDIDLLSAASKPIIARPEIHGAQPLDGV 282  
Db 164 VSNGRVN--GVLA--LSRALGDFFLK-----PYVIAEPDVTVE-LTEKD 203  
QY 283 GFLVLMSEGLYKALEAAHGPQANQETAAMIDTEFAKQTSLDVAQAQAVD 332  
Db 204 DFLILASDGLWDVL-----SNQEVVDIVRKHLKSGGPQEA-AKKLID 244

Search completed: December 22, 2004, 14:15:21  
Job time : 98.2825 secs

**This Page Blank (uspio)**

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 13:58:53 ; Search time 28.8698 Seconds  
(without alignments)  
1330.044 Million cell updates/sec

Title: US-09-830-144-2  
Perfect score: 3014  
Sequence: 1 MSTASAASSSSSSAGEMIE.....QCKKQLEVIRSQQKRGQTS 579

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3014	100.0	579	4	US-09-529-279-4
2	3014	100.0	579	4	US-10-158-895-4
3	3014	100.0	590	4	US-09-529-279-15
4	3014	100.0	590	4	US-10-158-895-15
5	481	16.0	455	3	US-09-221-235-5
6	481	16.0	455	3	US-09-221-928-5
7	481	16.0	455	3	US-09-221-527-5
8	481	16.0	455	3	US-09-221-236-5
9	481	16.0	455	3	US-09-221-416-5
10	481	16.0	455	3	US-09-221-245-5
11	481	16.0	455	3	US-09-163-115-5
12	481	16.0	455	3	US-09-221-528-5
13	481	16.0	455	3	US-09-593-553-5
14	481	16.0	455	3	US-09-221-237-5
15	481	16.0	455	3	US-09-399-588-2
16	481	16.0	455	4	US-09-757-982-5
17	461.5	15.3	1036	4	US-10-014-882-2
18	418.5	13.9	394	4	US-09-345-473E-19
19	418	13.9	835	3	US-09-291-839-2
20	418	13.9	835	4	US-09-458-457-2
21	418	13.9	835	4	US-09-947-199A-2
22	416	13.8	328	4	US-09-345-473E-18
23	414	13.7	835	4	US-09-458-457-8
24	414	13.7	835	4	US-09-947-199A-8
25	412	13.7	668	1	US-08-205-018-2
26	412	13.7	859	1	US-08-395-580-2
27	412	13.7	859	5	PCT-US95-02792-2

28	398	13.2	821	1	US-07-928-464-2	Sequence 2, Appli
29	398	13.2	821	1	US-08-003-311B-2	Sequence 2, Appli
30	398	13.2	821	1	US-08-261-432-2	Sequence 2, Appli
31	398	13.2	821	5	PCT-US93-07347-2	Sequence 2, Appli
32	393	13.0	263	3	US-09-035-706-5	Sequence 5, Appli
33	393	13.0	263	3	US-08-955-841-5	Sequence 5, Appli
34	393	13.0	263	3	US-09-390-425-5	Sequence 5, Appli
35	393	13.0	263	3	US-09-566-906-5	Sequence 5, Appli
36	393	13.0	263	4	US-09-589-034-5	Sequence 5, Appli
37	392	13.0	850	4	US-09-904-389-2	Sequence 2, Appli
38	390.5	13.0	590	4	US-09-312-283C-409	Sequence 409, App
39	390.5	13.0	786	4	US-09-509-802-2	Sequence 2, Appli
40	390.5	13.0	787	3	US-09-188-930-334	Sequence 334, App
41	390.5	13.0	787	4	US-09-312-283C-334	Sequence 334, App
42	380.5	12.6	536	3	US-09-188-930-185	Sequence 185, App
43	380.5	12.6	536	4	US-09-312-283C-185	Sequence 185, App
44	379.5	12.6	1584	3	US-09-457-040B-27	Sequence 27, Appli
45	375	12.4	784	4	US-09-781-882-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-529-279-4  
; Sequence 4, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-4

Query Match	100.0%	Score	3014;	DB	4;	Length	579;
Best Local Similarity	100.0%	Pred. No.	3.6e-222;				
Matches	579;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MSTASAASSSSSSAGEMIEAPSOVLNFEEDYKIEIEVEEVVGRGAFGVCKAKRAKDV	60				
Db	1	MSTASAASSSSSSSSAGEMIEAPSOVLNFEEDYKIEIEVEEVVGRGAFGVCKAKRAKDV	60				
QY	61	AIKQIESERKAFIVELRQLSRVNHNPVIXLYGACLNPCVLMVEYAEGLSYNLVHGA	120				
Db	61	AIKQIESERKAFIVELRQLSRVNHNPVIXLYGACLNPCVLMVEYAEGLSYNLVHGA	120				
QY	121	PLPYTTAAHMSWCLQCSQGVYALHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC	180				
Db	121	PLPYTTAAHMSWCLQCSQGVYALHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC	180				
QY	181	DIQTHMTNNKGAAMAPVEFEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAPRIM	240				
Db	181	DIQTHMTNNKGAAMAPVEFEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAPRIM	240				
QY	241	WAVNGTRPPLIKNLKPPIESLMTRCWSKDSQPSMBEIVKIMTHLMRYFPFGADEPQY	300				
Db	241	WAVNGTRPPLIKNLKPPIESLMTRCWSKDSQPSMBEIVKIMTHLMRYFPFGADEPQY	300				
QY	301	PCOYSDSGQNSATSTGSMFDIASNTNSKSDTNMEQVPATNDTIKRLSKLLKNQAKQ	360				

Db 301 PCQYSDGQNSATSTGSPMDIANTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360  
Qy 361 SESGRSLGASHGSSVESLPPTSEKGRMSADMSIEARIAATNGQPRRSIQDLTVTG 420  
Db 361 SESGRSLGASHGSSVESLPPTSEKGRMSADMSIEARIAATNGQPRRSIQDLTVTG 420  
Qy 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480  
Db 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480  
Qy 481 QPLAPCNSKESMAVFEQCHKMAQYMKVQTEIALLQKQELVAELDDQDKQONTSL 540  
Db 481 QPLAPCNSKESMAVFEQCHKMAQYMKVQTEIALLQKQELVAELDDQDKQONTSL 540  
Qy 541 VOEHKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579  
Db 541 VOEHKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579

## RESULT 2

US-10-158-895-4  
; Sequence 4, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-895-4

Query Match 100.0%; Score 3014; DB 4; Length 579;  
Best Local Similarity 100.0%; Pred. No. 3.6e-222;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSTAASASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKRAKDV 60  
Db 1 MSTAASASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKRAKDV 60  
Qy 61 AIKQIESSEKAFIVELRQLSRVNHNPVILKLYGACLNPCVLVMEYAEAGGSLYNVLHGA 120  
Db 61 AIKQIESSEKAFIVELRQLSRVNHNPVILKLYGACLNPCVLVMEYAEAGGSLYNVLHGA 120  
Qy 121 PLPYVTAHAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTAC 180  
Db 121 PLPYVTAHAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTAC 180  
Qy 181 DIQHTMTNNKGSAAWAPVEFGSNYSKCDVFSWGIIILWEVITRKPDEIGGPAFRIM 240  
Db 181 DIQHTMTNNKGSAAWAPVEFGSNYSKCDVFSWGIIILWEVITRKPDEIGGPAFRIM 240  
Qy 241 WAVHNGTRPPLIKNPKPIESLMTKCSKDPQSPQSMIEIVKIMTHLMRYFFGADPELOY 300  
Db 241 WAVHNGTRPPLIKNPKPIESLMTKCSKDPQSPQSMIEIVKIMTHLMRYFFGADPELOY 300  
Qy 301 PCQYSDGQNSATSTGSPMDIANTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360  
Db 301 PCQYSDGQNSATSTGSPMDIANTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360

Qy 361 SESGRSLGASHGSSVESLPPTSEKGRMSADMSIEARIAATNGQPRRSIQDLTVTG 420  
Db 361 SESGRSLGASHGSSVESLPPTSEKGRMSADMSIEARIAATNGQPRRSIQDLTVTG 420  
Qy 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480  
Db 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480  
Qy 481 QPLAPCNSKESMAVFEQCHKMAQYMKVQTEIALLQKQELVAELDDQDKQONTSL 540  
Db 481 QPLAPCNSKESMAVFEQCHKMAQYMKVQTEIALLQKQELVAELDDQDKQONTSL 540  
Qy 541 VOEHKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579  
Db 541 VOEHKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579

## RESULT 3

US-09-529-279-15  
; Sequence 15, Application US/09529279  
; Patent No. 6451617  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/09/529,279  
; CURRENT FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-529-279-15

Query Match 100.0%; Score 3014; DB 4; Length 590;  
Best Local Similarity 100.0%; Pred. No. 3.7e-222;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSTAASASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKRAKDV 60  
Db 1 MSTAASASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKRAKDV 60  
Qy 61 AIKQIESSEKAFIVELRQLSRVNHNPVILKLYGACLNPCVLVMEYAEAGGSLYNVLHGA 120  
Db 61 AIKQIESSEKAFIVELRQLSRVNHNPVILKLYGACLNPCVLVMEYAEAGGSLYNVLHGA 120  
Qy 121 PLPYVTAHAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTAC 180  
Db 121 PLPYVTAHAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTAC 180  
Qy 181 DIQHTMTNNKGSAAWAPVEFGSNYSKCDVFSWGIIILWEVITRKPDEIGGPAFRIM 240  
Db 181 DIQHTMTNNKGSAAWAPVEFGSNYSKCDVFSWGIIILWEVITRKPDEIGGPAFRIM 240  
Qy 241 WAVHNGTRPPLIKNPKPIESLMTKCSKDPQSPQSMIEIVKIMTHLMRYFFGADPELOY 300  
Db 241 WAVHNGTRPPLIKNPKPIESLMTKCSKDPQSPQSMIEIVKIMTHLMRYFFGADPELOY 300  
Qy 301 PCQYSDGQNSATSTGSPMDIANTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360  
Db 301 PCQYSDGQNSATSTGSPMDIANTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360  
Qy 361 SESGRSLGASHGSSVESLPPTSEKGRMSADMSIEARIAATNGQPRRSIQDLTVTG 420  
Db 361 SESGRSLGASHGSSVESLPPTSEKGRMSADMSIEARIAATNGQPRRSIQDLTVTG 420



QY 421 TEPGVSSRSSSPSVRMITTSKPTSEKPTRSHPTDDSTDTNGSDNSIPMAYLTLDHQL 480  
Db 421 TEPGVSSRSSSPSVRMITTSKPTSEKPTRSHPTDDSTDTNGSDNSIPMAYLTLDHQL 480  
QY 481 QPLAPCNSKESMAVFEQHCMAQYMKVQTEIALLQKQELVAELDDQDEKQOQNTSRL 540  
Db 481 QPLAPCNSKESMAVFEQHCMAQYMKVQTEIALLQKQELVAELDDQDEKQOQNTSRL 540  
QY 541 VQEHKKLLDENKSLSTYYQCKKQLEVIQSQQKROGTS 579  
Db 541 VQEHKKLLDENKSLSTYYQCKKQLEVIQSQQKROGTS 579

## RESULT 4

US-10-158-895-15  
; Sequence 15, Application US/10158895  
; Patent No. 6551840  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-895-15

Query Match 100.0%; Score 3014; DB 4; Length 590;  
Best Local Similarity 100.0%; Pred. No. 3,7e-222;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSTAASASSSSSSAGEMEATPSQVLNFEIDYKEIIEVEEVGRCGAFVGVCKAKWRAKDV 60  
Db 1 MSTAASASSSSSSAGEMEATPSQVLNFEIDYKEIIEVEEVGRCGAFVGVCKAKWRAKDV 60  
QY 61 AIKQIESESERKAFIVELRQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGSLYVNLHGAE 120  
Db 61 AIKQIESESERKAFIVELRQLSRVNHNPVIVKLYGACLPVCLVMEYAEAGSLYVNLHGAE 120  
QY 121 PLPYVYTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGVVLKICDPGTAC 180  
Db 121 PLPYVYTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGVVLKICDPGTAC 180  
QY 181 DIQHTMTNNKGAAMAPVFEFGSNVSEKCDVFSGIIILWEVITRRKPPDETGGAFRIM 240  
Db 181 DIQHTMTNNKGAAMAPVFEFGSNVSEKCDVFSGIIILWEVITRRKPPDETGGAFRIM 240  
QY 241 WAVHNGTRPPLKNIKPPIESLMTRCWSKDPQSPSMEEIVKIMTHLMRYFFGADEPLOY 300  
Db 241 WAVHNGTRPPLKNIKPPIESLMTRCWSKDPQSPSMEEIVKIMTHLMRYFFGADEPLOY 300  
QY 301 PCQYDEGQSNATSGSPMDIASNNTSKSDTNMEQVPAVNDTIKRLSKLLKNQAKQ 360  
Db 301 PCQYDEGQSNATSGSPMDIASNNTSKSDTNMEQVPAVNDTIKRLSKLLKNQAKQ 360  
QY 361 SESGRSLGASHGSSVESLPPTSEKGRMSADMSIEARIAATTGNGQPRRSIQDLTVTG 420  
Db 361 SESGRSLGASHGSSVESLPPTSEKGRMSADMSIEARIAATTGNGQPRRSIQDLTVTG 420  
QY 421 TEPGVSSRSSSPSVRMITTSKPTSEKPTRSHPTDDSTDTNGSDNSIPMAYLTLDHQL 480

Db 421 TEPGVSSRSSSPSVRMITTSKPTSEKPTRSHPTDDSTDTNGSDNSIPMAYLTLDHQL 480  
QY 481 QPLAPCNSKESMAVFEQHCMAQYMKVQTEIALLQKQELVAELDDQDEKQOQNTSRL 540  
Db 481 QPLAPCNSKESMAVFEQHCMAQYMKVQTEIALLQKQELVAELDDQDEKQOQNTSRL 540  
QY 541 VQEHKKLLDENKSLSTYYQCKKQLEVIQSQQKROGTS 579  
Db 541 VQEHKKLLDENKSLSTYYQCKKQLEVIQSQQKROGTS 579

## RESULT 5

US-09-221-235-5  
; Sequence 5, Application US/09221235  
; Patent No. 6043040  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,235  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-235-5

Query Match 16.0%; Score 481; DB 3; Length 455;  
Best Local Similarity 30.4%; Pred. No. 1.1e-28;  
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;  
QY 27 NFEEDYKEIIEVEEVGRCGAFVGVCKAKW-RAKQVAIK---QIESESERKAFIVELRQL 81  
Db 7 SFVQIKFDLQPFENGCGGSGFVYRAKVISQDKEVAVKKLIKIEKEAE-----IL 57  
QY 82 SRVNHNPVIVKLYGACLP--VCLVMEYAEAGSLYVNLHG--AEPILPYVYTAAHAMSWCLQ 137  
Db 58 SVLSHRNIIQFVGLVLEPNYGVIVTEYASLSLYDYNRSEEM---DMDHMTWATDV 114  
QY 138 SQGVAYLHSMQPKALIHRLDKPNNLLVAGGVVLKICDPGTACDIQHTMTNNK--GSA 195  
Db 115 AKGMHYLHMEAPVKVTHRDLKSRNVVIAADG-VLKICDPG-ASRPHNHTTHMSLVGTF 172  
QY 196 MAPEVFEFGSNVSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMW-AVHNGTRPPLK 254  
Db 173 MAPEVIOQLPVSSETCDTYSYGVVLWEMLTREVPFKLEG--LQVAVLVVEKNERLTIPS 230  
QY 255 LPKPIESLMTRCWSKDPQSPSMEEIVKIMTHLMRYFFGADEPLOYCQYSDGOSNSAT 314  
Db 231 CPRSFAELLHQWEADAKRPSFKIISL-----ESNSNDT- 267  
QY 315 STGSPMDIASNNTSKSDTNMEQVPAVNDTIKRLSKLLKNQAKQCSBGRSL---GAS 371  
Db 268 ---SLPDKCNSPLHNKAEWRCE-IEATLERLKLRLDLKRLKRLKRLKRLKRLKRLK 323  
QY 372 HGSSVESLPPTS-----EGKMSADMSIEARIAATTGNGQPRRSIQDLTVTG 421  
Db 324 QOSNTPLLLPAAARMSSESYFESKTESAESMSQITATSNNGEGHGMNPSLQAMLMGF 383  
QY 422 EPGQVSSRSSSPSV 435  
Db 384 --GDIFSNNKAGAV 395

## RESULT 6

US-09-221-928-5  
; Sequence 5, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:





```
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

Query Match      16.0%; Score 481; DB 3; Length 455;
Best Local Similarity 30.4%; Pred. No. 1.1e-28;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEIDYKEIEVEEVGAGFVCKAKW--RAKDVAIK---QIESEERKAFIVELQOL 81
DB 7 SFVQIKFDLQFFENCNGGSGFVYRAKWSQDEKAVKLLKIEKAE-----IL 57

QY 82 SRVNHNTVKLYGACLN--VCLVMEYAEAGSLYNVLHG--AEPLPYTAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSRSEEM---DMDHMTWATDV 114

QY 138 SQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHTMNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNETTHMSLVGTFPW 172

QY 196 MAPEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLEG--LQVAWLVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPQSRPSMBEIVKIMTHLMRYFPFGADEPLQPCQYSDGOSNAT 314
DB 231 CPRFAELHQCWEADAKKRFKQIISIL-----ESMSNDT- 267

QY 315 STGSFMDIASNTNKNSTNNMEQVPATNDTKRLSKLLKNQAKQOSSEGRSL----GAS 371
DB 268 ---SLPDKCNFLNKAERCE-IEATLERLUKLERDLSFKEQELKERERLKNWEQKLT 323

QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
DB 324 EQSNTPLLPLAARMSSESYFESKTESNSAEMSCQITATNGEGHGMNPSIQAMLMWGF 383

QY 422 EPGQVSSRSSSPSV 435
DB 384 --GDIFSNNKAGAV 395

RESULT 13
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

Query Match      16.0%; Score 481; DB 3; Length 455;
Best Local Similarity 30.4%; Pred. No. 1.1e-28;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEIDYKEIEVEEVGAGFVCKAKW--RAKDVAIK---QIESEERKAFIVELQOL 81
DB 7 SFVQIKFDLQFFENCNGGSGFVYRAKWSQDEKAVKLLKIEKAE-----IL 57

QY 82 SRVNHNTVKLYGACLN--VCLVMEYAEAGSLYNVLHG--AEPLPYTAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSRSEEM---DMDHMTWATDV 114

QY 138 SQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHTMNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNETTHMSLVGTFPW 172

QY 196 MAPEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLEG--LQVAWLVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPQSRPSMBEIVKIMTHLMRYFPFGADEPLQPCQYSDGOSNAT 314
DB 231 CPRFAELHQCWEADAKKRFKQIISIL-----ESMSNDT- 267

QY 315 STGSFMDIASNTNKNSTNNMEQVPATNDTKRLSKLLKNQAKQOSSEGRSL----GAS 371
DB 268 ---SLPDKCNFLNKAERCE-IEATLERLUKLERDLSFKEQELKERERLKNWEQKLT 323

QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
DB 324 EQSNTPLLPLAARMSSESYFESKTESNSAEMSCQITATNGEGHGMNPSIQAMLMWGF 383

QY 422 EPGQVSSRSSSPSV 435
DB 384 --GDIFSNNKAGAV 395

RESULT 12
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-528-5

Query Match      16.0%; Score 481; DB 3; Length 455;
Best Local Similarity 30.4%; Pred. No. 1.1e-28;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEIDYKEIEVEEVGAGFVCKAKW--RAKDVAIK---QIESEERKAFIVELQOL 81
DB 7 SFVQIKFDLQFFENCNGGSGFVYRAKWSQDEKAVKLLKIEKAE-----IL 57

QY 82 SRVNHNTVKLYGACLN--VCLVMEYAEAGSLYNVLHG--AEPLPYTAHAMSWCLQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLDYINSRSEEM---DMDHMTWATDV 114

QY 138 SQGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHTMNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNETTHMSLVGTFPW 172

QY 196 MAPEVFEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVLMWMLTREVPFKGLEG--LQVAWLVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPQSRPSMBEIVKIMTHLMRYFPFGADEPLQPCQYSDGOSNAT 314
DB 231 CPRFAELHQCWEADAKKRFKQIISIL-----ESMSNDT- 267

QY 315 STGSFMDIASNTNKNSTNNMEQVPATNDTKRLSKLLKNQAKQOSSEGRSL----GAS 371
DB 268 ---SLPDKCNFLNKAERCE-IEATLERLUKLERDLSFKEQELKERERLKNWEQKLT 323

QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
DB 324 EQSNTPLLPLAARMSSESYFESKTESNSAEMSCQITATNGEGHGMNPSIQAMLMWGF 383

QY 422 EPGQVSSRSSSPSV 435
DB 384 --GDIFSNNKAGAV 395
```

Db 268 ---SLPDKCNSFLHKAWEKCE-IEATLERLKKLRLDLSFKQEQLKERERLKKWMEQKLT 323  
QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421  
Db 324 EQSNTPLLLPLAARMSEESYFESKTEESNAEMSCQITATSNGEHGMPNPSLQAMLMGF 383  
QY 422 EPGQVSSRSSSPSV 435  
Db 384 --GDIFSMNKAGAV 395

## RESULT 14

US-09-221-237-5  
; Sequence 5, Application US/09221237  
; Patent No. 6214597  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,237  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-237-5

Query Match 16.0%; Score 481; DB 3; Length 455;  
Best Local Similarity 30.4%; Pred. No. 1.le-28;  
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;  
QY 27 NFEEDYKEIEVEEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81  
Db 7 SFVQIKFDLQFFENCGGSGFSGVYRAKWSQDKEVAVKLLKIEKAE-----IL 57  
QY 82 SRVNHPIVKLYGACLN--VCLVMEYAEGLSYNVLHG--AEPLPYTAAHMSWCLQC 137  
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYINSNRSEEM--DMDHMTWATDV 114  
QY 138 SOGVAYLHSMQKALIHRLDKPPNLLVAGTVLKLCDFTACTDQTHMTNNK--GSAAW 195  
Db 115 AKGMHYLHMEAPVKVTHRLKSRNVVIAADG-VLKICDFG-ASRFHNHTHMSLVGTFPW 172  
QY 196 MAPEVEGSGNYSEKCDVFSWGIILWEVITRRKPFQIEIGGPAFRIMW-AVHNGTRPPLIKN 254  
Db 173 MAPEVIQSLPVSETCDTYSYGVVLEMLTREVPFKLEG--LQVAVLVVEKNERLTIPS 230  
QY 255 LPKPTESLMTRCWSKDPQORPMBEIVKIMTHLMRYFPGADEPLQYPCQYSDGQSNAT 314  
Db 231 CPRSFAELLHQWEADAKKRPSEFKQIISIL-----ESMSNDT- 267  
QY 315 STGSPMDIATNTSNKSDTNMEQVATNDTIKRLSKLLKNQAKQOSESGRSL---GAS 371  
Db 268 ---SLPDKCNSFLHKAWEKCE-IEATLERLKKLRLDLSFKQEQLKERERLKKWMEQKLT 323  
QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421  
Db 324 EQSNTPLLLPLAARMSEESYFESKTEESNAEMSCQITATSNGEHGMPNPSLQAMLMGF 383  
QY 422 EPGQVSSRSSSPSV 435  
Db 384 --GDIFSMNKAGAV 395

## RESULT 15

US-09-399-588-2  
; Sequence 2, Application US/09399588  
; Patent No. 6511825  
; GENERAL INFORMATION:

; APPLICANT: Ruggieri, Rosamaria  
; APPLICANT: Callow, Marinella  
; APPLICANT: Diaz, Paul W.  
; TITLE OF INVENTION: No. 6511825el Cell Signaling Polypeptides and Nucleic Acids  
; FILE REFERENCE: 1044-US  
; CURRENT APPLICATION NUMBER: US/09/399,588  
; CURRENT FILING DATE: 1999-09-20  
; EARLIER APPLICATION NUMBER: 60/104,088  
; EARLIER FILING DATE: 1998-10-13  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Human J42  
US-09-399-588-2

Query Match 16.0%; Score 481; DB 4; Length 455;  
Best Local Similarity 30.4%; Pred. No. 1.le-28;  
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;  
QY 27 NFEEDYKEIEVEEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELROL 81  
Db 7 SFVQIKFDLQFFENCGGSGFSGVYRAKWSQDKEVAVKLLKIEKAE-----IL 57  
QY 82 SRVNHPIVKLYGACLN--VCLVMEYAEGLSYNVLHG--AEPLPYTAAHMSWCLQC 137  
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYINSNRSEEM--DMDHMTWATDV 114  
QY 138 SOGVAYLHSMQKALIHRLDKPPNLLVAGTVLKLCDFTACTDQTHMTNNK--GSAAW 195  
Db 115 AKGMHYLHMEAPVKVTHRLKSRNVVIAADG-VLKICDFG-ASRFHNHTHMSLVGTFPW 172  
QY 196 MAPEVEGSGNYSEKCDVFSWGIILWEVITRRKPFQIEIGGPAFRIMW-AVHNGTRPPLIKN 254  
Db 173 MAPEVIQSLPVSETCDTYSYGVVLEMLTREVPFKLEG--LQVAVLVVEKNERLTIPS 230  
QY 255 LPKPTESLMTRCWSKDPQORPMBEIVKIMTHLMRYFPGADEPLQYPCQYSDGQSNAT 314  
Db 231 CPRSFAELLHQWEADAKKRPSEFKQIISIL-----ESMSNDT- 267  
QY 315 STGSPMDIATNTSNKSDTNMEQVATNDTIKRLSKLLKNQAKQOSESGRSL---GAS 371  
Db 268 ---SLPDKCNSFLHKAWEKCE-IEATLERLKKLRLDLSFKQEQLKERERLKKWMEQKLT 323  
QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421  
Db 324 EQSNTPLLLPLAARMSEESYFESKTEESNAEMSCQITATSNGEHGMPNPSLQAMLMGF 383  
QY 422 EPGQVSSRSSSPSV 435  
Db 384 --GDIFSMNKAGAV 395

Search completed: December 22, 2004, 13:59:57  
Job time : 30.8698 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 13:58:54 ; Search time 103.717 Seconds  
(without alignments)  
1997.604 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSSAGEMIE.....QCKKQLEVIRSOQKXRGTS 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1599859 seqs, 357834939 residues

Total number of hits satisfying chosen parameters: 1599859

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	579	13	US-10-158-895-4
2	3014	100.0	579	14	US-10-364-743-4
3	3014	100.0	590	13	US-10-158-895-4
4	3014	100.0	590	14	US-10-364-743-15
5	2488.5	82.6	518	14	US-10-283-023-2
6	2488.5	82.6	518	15	US-10-386-414-13
7	1781	59.1	336	16	US-10-664-421-135
8	502	16.7	953	14	US-10-369-022-56
9	502	16.7	953	15	US-10-042-865-98
10	493.5	16.4	954	15	US-10-042-865-97
11	493.5	16.4	954	15	US-10-263-929-187
12	482.5	16.0	330	16	US-10-664-421-136
13	482	16.0	473	15	US-10-296-115-837

14	481	16.0	455	9	US-09-757-982-5	Sequence 5, Appli
15	481	16.0	455	14	US-10-094-749-2477	Sequence 2477, Ap
16	481	16.0	455	15	US-10-352-674A-2	Sequence 2, Appli
17	462.5	15.3	1066	15	US-10-042-865-95	Sequence 95, Appl
18	462.5	15.3	1096	15	US-10-210-130-130	Sequence 130, App
19	462.5	15.3	1118	15	US-10-112-944-259	Sequence 259, App
20	462	15.3	1021	15	US-10-451-168-87	Sequence 87, Appl
21	461.5	15.3	719	15	US-10-182-243-55	Sequence 55, Appl
22	461.5	15.3	850	15	US-10-263-929-192	Sequence 192, App
23	461.5	15.3	1036	13	US-10-014-882-2	Sequence 2, Appli
24	461.5	15.3	1036	14	US-10-354-358-24	Sequence 24, Appli
25	461.5	15.3	1036	14	US-10-419-279-2	Sequence 2, Appli
26	461.5	15.3	1036	15	US-10-263-929-189	Sequence 189, App
27	461.5	15.3	1036	16	US-10-473-670-9	Sequence 9, Appli
28	461.5	15.3	1036	17	US-10-478-068-2	Sequence 2, Appli
29	461.5	15.3	1036	17	US-10-618-941-118	Sequence 118, App
30	459.5	15.2	847	13	US-10-143-133-2	Sequence 2, Appli
31	459.5	15.2	847	15	US-10-263-929-188	Sequence 188, App
32	459.5	15.2	847	15	US-10-380-235-2	Sequence 2, Appli
33	456.5	15.1	800	16	US-10-408-765A-1101	Sequence 1101, Ap
34	456.5	15.1	800	17	US-10-737-450-66	Sequence 66, Appl
35	456.5	15.1	1036	16	US-10-408-765A-2675	Sequence 2675, Ap
36	456	15.1	948	13	US-10-087-192-1869	Sequence 1869, Ap
37	455	15.1	312	16	US-10-664-421-137	Sequence 137, App
38	454	15.1	964	15	US-10-210-130-32	Sequence 32, Appl
39	451.5	15.0	1097	14	US-10-288-798-12	Sequence 12, Appl
40	451.5	15.0	1097	15	US-10-362-892-12	Sequence 12, Appl
41	451	15.0	1046	15	US-10-258-106-9	Sequence 9, Appli
42	450.5	14.9	1002	15	US-10-263-929-193	Sequence 193, App
43	445.5	14.8	376	15	US-10-424-599-271998	Sequence 271998, App
44	444.5	14.7	1024	15	US-10-042-865-16	Sequence 16, Appl
45	439.5	14.6	422	16	US-10-437-963-132604	Sequence 132604,

#### ALIGNMENTS

RESULT 1  
US-10-158-895-4  
; Sequence 4, Application US/10158895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JF98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-895-4

Query Match 100.0%; Score 3014; DB 13; Length 579;  
Best Local Similarity 100.0%; Pred. No. 3.6e-193;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSTAASASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGWCKAKWRADV 60  
Db 1 MSTAASASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGWCKAKWRADV 60  
Qy 61 AIKQIESESEKARTVELRQLSRVNHNVKLYGACLNFCVLMVEYABGSSLYNVHGAEE 120

Db 61 AIKQIESERKAFIVELRQLSRVNHNPVVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120  
Qy 121 PLPYTAAHAMSCLQCSQGVAYLHSMQPKALHRDLKPPNLLVAGTGLVKICDFGTAC 180  
Db 121 PLPYTAAHAMSCLQCSQGVAYLHSMQPKALHRDLKPPNLLVAGTGLVKICDFGTAC 180  
Qy 181 DIQTHMTNNKGSAAWMAPEVFEAGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
Db 181 DIQTHMTNNKGSAAWMAPEVFEAGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
Qy 241 WAVINGTRPPLIKNLPKPIESLMTRCWSKOPSPQSPSMBEIVKIMTHLMRYPPGGADEPLQY 300  
Db 241 WAVINGTRPPLIKNLPKPIESLMTRCWSKOPSPQSPSMBEIVKIMTHLMRYPPGGADEPLQY 300  
Qy 301 PCQYSDGQNSATSTGSMFMDIASNTSNKSDTNEQVPAINDTIKRLSKLLKNQAKQ 360  
Db 301 PCQYSDGQNSATSTGSMFMDIASNTSNKSDTNEQVPAINDTIKRLSKLLKNQAKQ 360  
Qy 361 SESGRSLGASHGSSVESLPPTSGKMSADMSIEARIAATTGNGOPRRRSIIDLTVTG 420  
Db 361 SESGRSLGASHGSSVESLPPTSGKMSADMSIEARIAATTGNGOPRRRSIIDLTVTG 420  
Qy 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPTPDDSTDTNGSDNSIPMAYLTLDHQL 480  
Db 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPTPDDSTDTNGSDNSIPMAYLTLDHQL 480  
Qy 481 QPLAPCNSKESMAVFEHQCKMAQBYMKVQTEIALLLQKQELVAELDDQDKDQNTSRL 540  
Db 481 QPLAPCNSKESMAVFEHQCKMAQBYMKVQTEIALLLQKQELVAELDDQDKDQNTSRL 540  
Qy 541 VQEHKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579  
Db 541 VQEHKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579

## RESULT 2

US-10-384-743-4  
; Sequence 4, Application US/10384743  
; Publication No. US20030162228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 579  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-384-743-4

Query Match 100.0%; Score 3014; DB 14; Length 579;  
Best Local Similarity 100.0%; Pred. No. 3,6e-193;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSTAASASSSSSSSAGEMIEAPSVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60  
Db 1 MSTAASASSSSSSSAGEMIEAPSVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60  
Qy 61 AIKQIESERKAFIVELRQLSRVNHNPVVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120  
Db 61 AIKQIESERKAFIVELRQLSRVNHNPVVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120

Qy 121 PLPYTAAHAMSCLQCSQGVAYLHSMQPKALHRDLKPPNLLVAGTGLVKICDFGTAC 180  
Db 121 PLPYTAAHAMSCLQCSQGVAYLHSMQPKALHRDLKPPNLLVAGTGLVKICDFGTAC 180  
Qy 181 DIQTHMTNNKGSAAWMAPEVFEAGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
Db 181 DIQTHMTNNKGSAAWMAPEVFEAGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
Qy 241 WAVINGTRPPLIKNLPKPIESLMTRCWSKOPSPQSPSMBEIVKIMTHLMRYPPGGADEPLQY 300  
Db 241 WAVINGTRPPLIKNLPKPIESLMTRCWSKOPSPQSPSMBEIVKIMTHLMRYPPGGADEPLQY 300  
Qy 301 PCQYSDGQNSATSTGSMFMDIASNTSNKSDTNEQVPAINDTIKRLSKLLKNQAKQ 360  
Db 301 PCQYSDGQNSATSTGSMFMDIASNTSNKSDTNEQVPAINDTIKRLSKLLKNQAKQ 360  
Qy 361 SESGRSLGASHGSSVESLPPTSGKMSADMSIEARIAATTGNGOPRRRSIIDLTVTG 420  
Db 361 SESGRSLGASHGSSVESLPPTSGKMSADMSIEARIAATTGNGOPRRRSIIDLTVTG 420  
Qy 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPTPDDSTDTNGSDNSIPMAYLTLDHQL 480  
Db 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPTPDDSTDTNGSDNSIPMAYLTLDHQL 480  
Qy 481 QPLAPCNSKESMAVFEHQCKMAQBYMKVQTEIALLLQKQELVAELDDQDKDQNTSRL 540  
Db 481 QPLAPCNSKESMAVFEHQCKMAQBYMKVQTEIALLLQKQELVAELDDQDKDQNTSRL 540  
Qy 541 VQEHKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579  
Db 541 VQEHKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579

## RESULT 3

US-10-158-895-15  
; Sequence 15, Application US/10158895  
; Publication No. US20020155624A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/158,895  
; CURRENT FILING DATE: 2002-06-03  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-158-895-15

Query Match 100.0%; Score 3014; DB 13; Length 590;  
Best Local Similarity 100.0%; Pred. No. 3,7e-193;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSTAASASSSSSSSAGEMIEAPSVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60  
Db 1 MSTAASASSSSSSSAGEMIEAPSVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60  
Qy 61 AIKQIESERKAFIVELRQLSRVNHNPVVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120  
Db 61 AIKQIESERKAFIVELRQLSRVNHNPVVKLYGACLPVCLVMEYAEAGGSLYNVLHGAE 120  
Qy 121 PLPYTAAHAMSCLQCSQGVAYLHSMQPKALHRDLKPPNLLVAGTGLVKICDFGTAC 180



Db 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKCIDFGTAC 180  
QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
Db 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
QY 241 WAVHNGTRPPLIKNLPKPIESLMTWCWSDPSQPSMEEIVKIMTHLMRYFFGADPELOY 300  
Db 241 WAVHNGTRPPLIKNLPKPIESLMTWCWSDPSQPSMEEIVKIMTHLMRYFFGADPELOY 300  
QY 301 PCQYSDGQSNATSTGSPMDIASNTSNKSDTNMEQVPAATNDTIKRLSKLLKQAKQ 360  
Db 301 PCQYSDGQSNATSTGSPMDIASNTSNKSDTNMEQVPAATNDTIKRLSKLLKQAKQ 360  
QY 361 SSSGRLSLGASHGSSVESLPPTSEKGRMSADMSEIEARIAATTGNGQPPRRSIQDLTVTG 420  
Db 361 SSSGRLSLGASHGSSVESLPPTSEKGRMSADMSEIEARIAATTGNGQPPRRSIQDLTVTG 420  
QY 421 TEPGQVSSRSSSPSRVMTTSGTSEKPTRSHPTWPDSDTDNGSDNSIPMAYLTLHDOL 480  
Db 421 TEPGQVSSRSSSPSRVMTTSGTSEKPTRSHPTWPDSDTDNGSDNSIPMAYLTLHDOL 480  
QY 481 QPLAPCPNSKESMAVFEQHCMAQYMKVQTEIALLLQKQELVAELDDQDQNTSRL 540  
Db 481 QPLAPCPNSKESMAVFEQHCMAQYMKVQTEIALLLQKQELVAELDDQDQNTSRL 540  
QY 541 VOEHKKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579  
Db 541 VOEHKKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579

RESULT 4  
US-10-384-743-15  
; Sequence 15, Application US/10384743  
; Publication No. US20030162228A1  
; GENERAL INFORMATION:  
; APPLICANT: ONO, KOICHIRO  
; APPLICANT: OHTOMO, TOSHIHIKO  
; APPLICANT: TSUCHIYA, MASAYUKI  
; FILE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
; FILE REFERENCE: 053466/0278  
; CURRENT APPLICATION NUMBER: US/10/384,743  
; PRIOR FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: US/09/529,279  
; PRIOR FILING DATE: 2000-04-11  
; PRIOR APPLICATION NUMBER: PCT/JP98/04796  
; PRIOR FILING DATE: 1998-10-22  
; PRIOR APPLICATION NUMBER: JP 9/290188  
; PRIOR FILING DATE: 1997-10-22  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 590  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-384-743-15

Query Match 100.0%; Score 3014; DB 14; Length 590;  
Best Local Similarity 100.0%; Pred. No. 3.7e-193;  
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSTAAGSSSSSSSAGEMIEAPSVLNFEEDYKIEVEEVVGRGAFGVCKAKRAKDV 60  
Db 1 MSTAAGSSSSSSSAGEMIEAPSVLNFEEDYKIEVEEVVGRGAFGVCKAKRAKDV 60  
QY 61 AIKQIESSEKAFIVELRQLSRVNHNPVILKYGACLNPCVLWMEYAEAGSSLYNVLHGA 120  
Db 61 AIKQIESSEKAFIVELRQLSRVNHNPVILKYGACLNPCVLWMEYAEAGSSLYNVLHGA 120  
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKCIDFGTAC 180  
Db 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKCIDFGTAC 180

QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
Db 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
QY 241 WAVHNGTRPPLIKNLPKPIESLMTWCWSDPSQPSMEEIVKIMTHLMRYFFGADPELOY 300  
Db 241 WAVHNGTRPPLIKNLPKPIESLMTWCWSDPSQPSMEEIVKIMTHLMRYFFGADPELOY 300  
QY 301 PCQYSDGQSNATSTGSPMDIASNTSNKSDTNMEQVPAATNDTIKRLSKLLKQAKQ 360  
Db 301 PCQYSDGQSNATSTGSPMDIASNTSNKSDTNMEQVPAATNDTIKRLSKLLKQAKQ 360  
QY 361 SSSGRLSLGASHGSSVESLPPTSEKGRMSADMSEIEARIAATTGNGQPPRRSIQDLTVTG 420  
Db 361 SSSGRLSLGASHGSSVESLPPTSEKGRMSADMSEIEARIAATTGNGQPPRRSIQDLTVTG 420  
QY 421 TEPGQVSSRSSSPSRVMTTSGTSEKPTRSHPTWPDSDTDNGSDNSIPMAYLTLHDOL 480  
Db 421 TEPGQVSSRSSSPSRVMTTSGTSEKPTRSHPTWPDSDTDNGSDNSIPMAYLTLHDOL 480  
QY 481 QPLAPCPNSKESMAVFEQHCMAQYMKVQTEIALLLQKQELVAELDDQDQNTSRL 540  
Db 481 QPLAPCPNSKESMAVFEQHCMAQYMKVQTEIALLLQKQELVAELDDQDQNTSRL 540  
QY 541 VOEHKKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579  
Db 541 VOEHKKLLDENKSLSTYYQCKKQLEVRISQOQKRGTS 579

RESULT 5  
US-10-283-023-2  
; Sequence 2, Application US/10283023  
; Publication No. US20030091573A1  
; GENERAL INFORMATION:  
; APPLICANT: Garroll, Joseph M.  
; TITLE OF INVENTION: Methods and compositions for the  
; diagnosis and treatment of hematological disorders using  
; FILE OF INVENTION: 16319  
; FILE REFERENCE: MPI01-239P1RM  
; CURRENT APPLICATION NUMBER: US/10/283,023  
; PRIOR FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: 60/335,044  
; PRIOR FILING DATE: 2001-10-31  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-10-283-023-2

Query Match 82.6%; Score 2488.5; DB 14; Length 518;  
Best Local Similarity 93.8%; Pred. No. 4.4e-158;  
Matches 481; Conservative 0; Mismatches 5; Indels 27; Gaps 1;  
QY 1 MSTAAGSSSSSSSAGEMIEAPSVLNFEEDYKIEVEEVVGRGAFGVCKAKRAKDV 60  
Db 1 MSTAAGSSSSSSSAGEMIEAPSVLNFEEDYKIEVEEVVGRGAFGVCKAKRAKDV 60  
QY 61 AIKQIESSEKAFIVELRQLSRVNHNPVILKYGACLNPCVLWMEYAEAGSSLYNVLHGA 120  
Db 61 AIKQIESSEKAFIVELRQLSRVNHNPVILKYGACLNPCVLWMEYAEAGSSLYNVLHGA 120  
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKCIDFGTAC 180  
Db 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKCIDFGTAC 180  
QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
Db 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240  
QY 241 WAVHNGTRPPLIKNLPKPIESLMTWCWSDPSQPSMEEIVKIMTHLMRYFFGADPELOY 300  
Db 241 WAVHNGTRPPLIKNLPKPIESLMTWCWSDPSQPSMEEIVKIMTHLMRYFFGADPELOY 300

Db 241 WAVHNGTRPPLIKNLKPESLMTRCWSKDPQSPQSMEEIVKIMTHLMRYPPGADDEPLOY 300  
Qy 301 PCQYSDRGQNSATSTGSMFMDIATNTSNKSDTNNQVPAFNDTIKRLSKLLKNQAKQQ 360  
Db 301 PCQYSDRGQNSATSTGSMFMDIATNTSNKSDTNNQVPAFNDTIKRLSKLLKNQAKQQ 360  
Qy 361 SESGRSLGASHGSSVESLPTSTSEKRMADMSIEIARIAATT----- 403  
Db 361 SESGRSLGASHGSSVESLPTSTSEKRMADMSIEIARIAATTAYSKPKRGRHTASFGN 420  
Qy 404 -----GNGOPRRRSIQDLTVTGTGEPQVSSRSSSPSVRMITTSKPTSEKPTRSH 453  
Db 421 ILDVPEIVISGNGOPRRRSIQDLTVTGTGEPQVSSRSSSPSVRMITTSKPTSEKPTRSH 480  
Qy 454 WTPDDSTDTNGSDNSIPMAYILTDHQLQPLAPC 486  
Db 481 WTPDDSTDTNGSDNSIPMAYILTDHQLQARTSC 513

## RESULT 6

US-10-386-414-13  
; Sequence 13, Application US/10386414  
; Publication No. US2004006016A1  
; GENERAL INFORMATION:  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Robison, Keith E.  
; APPLICANT: White, David  
; APPLICANT: Williamson, Mark W.  
; APPLICANT: Cook, William James  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: MacBeth, Kyle J.  
; APPLICANT: Carroll, Joseph M.  
; APPLICANT: Chun, Miyoung  
; TITLE OF INVENTION: NOVEL 27875, 22025, 27420, 17906, 16319,  
; FILE OF INVENTION: 55092 AND 10218 MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MPI03-0210MNTM  
; CURRENT APPLICATION NUMBER: US/10/386,414  
; CURRENT FILING DATE: 2003-03-11  
; PRIOR APPLICATION NUMBER: 09/426,282  
; PRIOR FILING DATE: 1999-10-25  
; PRIOR APPLICATION NUMBER: 09/668,266  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 09/330,970  
; PRIOR FILING DATE: 1999-06-11  
; PRIOR APPLICATION NUMBER: 09/724,599  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/860,193  
; PRIOR FILING DATE: 2001-05-16  
; PRIOR APPLICATION NUMBER: 10/283,023  
; PRIOR FILING DATE: 2002-10-29  
; PRIOR APPLICATION NUMBER: 60/335,044  
; PRIOR FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: 10/010,943  
; PRIOR FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: 60/254,037  
; PRIOR FILING DATE: 2000-12-07  
; PRIOR APPLICATION NUMBER: 09/833,082  
; PRIOR FILING DATE: 2001-04-10  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 518  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-386-414-13

Query Match 82.6%; Score 2488.5; DB 15; Length 518;  
Best Local Similarity 93.8%; Pred. No. 4.4e-158;  
Matches 481; Conservative 0; Mismatches 5; Indels 27; Gaps 1;  
Qy 1 MSTAASASSSSSSAGEMIEAPSVLNFEIDYKIEIEVEVVGAFGVCKAKRAKDV 60  
Db 1 MSTAASASSSSSSAGEMIEAPSVLNFEIDYKIEIEVEVVGAFGVCKAKRAKDV 60

Qy 61 AIKOIESESEKAFIVELRQLSRVNHPIVKLYGACLNPCVLMVEYABGGSILYNVLHGAE 120  
Db 61 AIKOIESESEKAFIVELRQLSRVNHPIVKLYGACLNPCVLMVEYABGGSILYNVLHGAE 120  
Qy 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTGLKICDFGTAC 180  
Db 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTGLKICDFGTAC 180  
Qy 181 DIQHTMTNNKGAAMPAPEVPEGSNYSEKCDVFWGIIIMWVITRRKPFDBIGGPAFRIM 240  
Db 181 DIQHTMTNNKGAAMPAPEVPEGSNYSEKCDVFWGIIIMWVITRRKPFDBIGGPAFRIM 240  
Qy 241 WAVHNGTRPPLIKNLKPESLMTRCWSKDPQSPQSMEEIVKIMTHLMRYPPGADDEPLOY 300  
Db 241 WAVHNGTRPPLIKNLKPESLMTRCWSKDPQSPQSMEEIVKIMTHLMRYPPGADDEPLOY 300  
Qy 301 PCQYSDRGQNSATSTGSMFMDIATNTSNKSDTNNQVPAFNDTIKRLSKLLKNQAKQQ 360  
Db 301 PCQYSDRGQNSATSTGSMFMDIATNTSNKSDTNNQVPAFNDTIKRLSKLLKNQAKQQ 360  
Qy 361 SESGRSLGASHGSSVESLPTSTSEKRMADMSIEIARIAATT----- 403  
Db 361 SESGRSLGASHGSSVESLPTSTSEKRMADMSIEIARIAATTAYSKPKRGRHTASFGN 420  
Qy 404 -----GNGOPRRRSIQDLTVTGTGEPQVSSRSSSPSVRMITTSKPTSEKPTRSH 453  
Db 421 ILDVPEIVISGNGOPRRRSIQDLTVTGTGEPQVSSRSSSPSVRMITTSKPTSEKPTRSH 480  
Qy 454 WTPDDSTDTNGSDNSIPMAYILTDHQLQPLAPC 486  
Db 481 WTPDDSTDTNGSDNSIPMAYILTDHQLQARTSC 513

## RESULT 7

US-10-664-421-135  
; Sequence 135, Application US/10664421  
; Publication No. US20040142864A1  
; GENERAL INFORMATION:  
; APPLICANT: BREMER, RYAN  
; APPLICANT: IBRAHIM, PRABHA  
; APPLICANT: KUMAR, ABHINAV  
; APPLICANT: MANDIVAN, VALSAN  
; APPLICANT: MILBURN, MICHAEL V.  
; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE  
; FILE REFERENCE: 039363/0703  
; CURRENT APPLICATION NUMBER: US/10/664,421  
; CURRENT FILING DATE: 2003-09-16  
; PRIOR APPLICATION NUMBER: 60/412,341  
; PRIOR FILING DATE: 2002-09-20  
; PRIOR APPLICATION NUMBER: 60/411,398  
; PRIOR FILING DATE: 2002-09-16  
; NUMBER OF SEQ ID NOS: 169  
; SOFTWARE: Patent In Ver. 3.2  
; SEQ ID NO 135  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-664-421-135

Query Match 59.1%; Score 1781; DB 16; Length 336;  
Best Local Similarity 100.0%; Pred. No. 5.4e-111;  
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTAASASSSSSSAGEMIEAPSVLNFEIDYKIEIEVEVVGAFGVCKAKRAKDV 60  
Db 1 MSTAASASSSSSSAGEMIEAPSVLNFEIDYKIEIEVEVVGAFGVCKAKRAKDV 60  
Qy 61 AIKOIESESEKAFIVELRQLSRVNHPIVKLYGACLNPCVLMVEYABGGSILYNVLHGAE 120  
Db 61 AIKOIESESEKAFIVELRQLSRVNHPIVKLYGACLNPCVLMVEYABGGSILYNVLHGAE 120  
Qy 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTGLKICDFGTAC 180

```

Db 121 PLYPTTAHAMSQCQGVAYLHSMQPKALIHRLDKPNNLLVAGGVVLKICDFGTAC 180
QY 181 DIQHTMTNKGSAAMWAEVFEVGSNSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
Db 181 DIQHTMTNKGSAAMWAEVFEVGSNSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVINGTAPPLIKNPKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFFGADPELQY 300
Db 241 WAVINGTAPPLIKNPKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFFGADPELQY 300
QY 301 PCQYSDGQNSATSGFMDIASNTSNKSDTNME 336
Db 301 PCQYSDGQNSATSGFMDIASNTSNKSDTNME 336

RESULT 8
US-10-369-022-56
; Sequence 56, Application US/10369022
; Publication No. US20030203847A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Immaculada
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 9949, 14230, 760, 62553,
; TITLE OF INVENTION: 12216, 17719, 41897, 47174, 33408, 10002, 16209, 314, 636,
; TITLE OF INVENTION: 27410, 33260, 619, 15985, 69112, 2158, 224, 615, 44373,
; TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
; TITLE OF INVENTION: 13424 MOLECULES
; FILE REFERENCE: MPI02-027P1RNM01M
; CURRENT APPLICATION NUMBER: US/10/369,022
; PRIORITY FILING DATE: 2003-02-19
; PRIOR APPLICATION NUMBER: US 60/360,495
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/370,121
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/373,010
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: US 60/373,908
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/377,717
; PRIOR FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: US 60/379,949
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/382,409
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/385,280
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/386,879
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/387,536
; PRIOR FILING DATE: 2002-06-10
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-022-56

Query Match 16.7%; Score 502; DB 14; Length 953;
Best Local Similarity 29.3%; Pred. No. 7.2e-25;
Matches 147; Conservative 78; Mismatches 194; Indels 82; Gaps 15;

QY 21 APSQVNLNFEIDYKEIEVEVVGAGFVGVCKAKRPAKDVAIKQISESERKAFIV---- 76
Db 83 APAGLQLPQIEFHEILGIEITGVGGFGKVKYRALMRVEEVAVKAAARLDPEKDPAVTAQV 142
QY 77 --ELRQLSRVNHPTVKLYGACINP--VCLVMYEGGSLYNVLHGAEBPLPYTTAAHAMS 132
Db 143 CQEARLFGALQHPNIIALRGACINPHLCLVWEYARGGALSRLAGRRVPPHV-----LVN 198

```

```

QY 133 WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLV-----AGGTVLKICDFGTACD-IQT 184
Db 133 WCLQCSQGVAYLHSMQPKALIHRLDKPNNLLV-----AGGTVLKICDFGTACD-IQT 184
QY 199 WAVQVARGMNYLHNDAPVPIIHRDLKSNILILEAIENHNLAADTVLKITDFGLAREWHKT 258
Db 199 WAVQVARGMNYLHNDAPVPIIHRDLKSNILILEAIENHNLAADTVLKITDFGLAREWHKT 258
QY 185 HMTNKNKGSAAWMAPEVFEVGSNSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV- 243
Db 185 HMTNKNKGSAAWMAPEVFEVGSNSYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV- 243
QY 259 TKMSAAGTYANWMAPEVIRLSFSSKSDVMWFGVLLWELLTGEVVPYREI--DALAVAYGVA 316
Db 259 TKMSAAGTYANWMAPEVIRLSFSSKSDVMWFGVLLWELLTGEVVPYREI--DALAVAYGVA 316
QY 244 HNGTAPPLIKNPKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFFGADPELQYPCQ 303
Db 244 HNGTAPPLIKNPKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFFGADPELQYPCQ 303
QY 317 MNKLTLPSTCPEPFARLLBECWDPDGHRDFGSIILKRLVIEQ-----SALQMPLE 371
Db 317 MNKLTLPSTCPEPFARLLBECWDPDGHRDFGSIILKRLVIEQ-----SALQMPLE 371
QY 304 YSDEQNSATSGFMDIASNTSNKSDTNMEQVPAT-----NDTIRLESKILLKNO- 356
Db 304 YSDEQNSATSGFMDIASNTSNKSDTNMEQVPAT-----NDTIRLESKILLKNO- 356
QY 372 SFHSLQEDWKLEIQHMFDDLRKTKEKELRSREBELLAQAQORFQBEQLRRREQELAEREM 431
Db 372 SFHSLQEDWKLEIQHMFDDLRKTKEKELRSREBELLAQAQORFQBEQLRRREQELAEREM 431
QY 357 -----AKQOSESGRSLGASHGSSVESLP-----PTS 383
Db 357 -----AKQOSESGRSLGASHGSSVESLP-----PTS 383
QY 432 DIVERELHLLMCQLSQEKPRVKRKNFRAVLKUREGSSHLSLPSGFEHKITVQASPTL 491
Db 432 DIVERELHLLMCQLSQEKPRVKRKNFRAVLKUREGSSHLSLPSGFEHKITVQASPTL 491
QY 384 EGRKMSADMSEIARIAATNGQPPRRRSIQDLTVTGTGPGQVSSRSRSPSVRMITTSGP 443
Db 384 EGRKMSADMSEIARIAATNGQPPRRRSIQDLTVTGTGPGQVSSRSRSPSVRMITTSGP 443
QY 492 D-KRKGSDGASPPASPSII-----PRLRAIR--LTPVDCGSGSSGSSGSGTWSRGCP 542
Db 492 D-KRKGSDGASPPASPSII-----PRLRAIR--LTPVDCGSGSSGSSGSGTWSRGCP 542
QY 444 TSEK-----PTRSHWTPDDSD 459
Db 444 TSEK-----PTRSHWTPDDSD 459
QY 543 PKKEELVGGKKKGRTWGPSSST 563
Db 543 PKKEELVGGKKKGRTWGPSSST 563

RESULT 9
US-10-042-865-98
; Sequence 98, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10

```

```
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98
; LENGTH: 953
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-865-98

Query Match      16.7%; Score 502; DB 15; Length 953;
Best Local Similarity 29.3%; Pred. No. 7.2e-25;
Matches 147; Conservative 78; Mismatches 194; Indels 82; Gaps 15;

QY 21 APSQVLFEEIDYKEIEVEEVGAGVVCCKAKWRAKQVAIKQIESERKAFIV---- 76
Db 83 APAGLQLPQELPFHELOLEELIIGVGKGVTRALWEGEEVAVKARLDPEKDPVTAEOV 142
QY 77 --ELRQLSRVNHNPVIVKLYGACLPN--VCLVMVEAEGGSLYNVLHGAEPPLPYTTAAHAMS 132
Db 143 QCEARLFGALQHPNIIALRGACLPNPHLCLVMVEYARGGALSRLVLAGRRVPHV---LVN 198
QY 133 WCLQCSQGVAYLHSMQKALIHRLDKPPNLLV-----AGGTVLKICDFTGACD-IQT 184
Db 199 WAVQVARGMNYLHNDAPVPIIHRDLKSNIIILAEIENHNLDATVLKITDGLAREWHKT 258
QY 185 HMTNKGSAAMWAVEFEGSNYSKCDVFSNGIILWEVITRRKPFDEIGGPAFRIMWAV- 243
Db 259 TKMSAAGTYAMWAVEVIRLSFSKSSDWSVGVLLWELLTGEVYREI--DALAVAYGVA 316
QY 244 HNGTRPPLIKLPKPIESLMTRCKSKDPSQRPSEEEIVKIMTHLMRYPPGADPELOYPQ 303
Db 317 MNKLTLPSTCPEPFARLLDEECWDPDHPGRDPFGSILKRLVETQ-----SALFQMLE 371
QY 304 YSDGQNSATSTGSPMDIANTNTSNKSDTNWQVPAT-----NDITKRLSKLLKNQ- 356
Db 372 SFHSLQEDWKLEIQHMFDDLTTRKELRSEEEILLRAAQEQRFQELRRREQELAREM 431
QY 357 -----AKQSSGRLSLGASHGSSVESLP-----PTS 383
Db 432 DIVERELHLLMCQLSQEKPRVKRKKGNPKRAVLKLRGSSHSILSPSGFEHKITVQASPTL 491
QY 384 EGKMSADMSEIARIAATTGNGQFRRRSIQDLTVTGTGPGQVSSRSSSPSVRMITTSGP 443
Db 492 D-KRKGSDGASPPASPSII-----PRLRAIR--LTPVDCGSSGSSGSGGTSWRRGP 542
QY 444 TSEK-----PTRSHPTPDDS 459
Db 543 PKBELVGGKKKGRWTGSPST 563

RESULT 10
US-10-042-865-97
; Sequence 97, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
```

```
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 954
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-042-865-97

Query Match      16.4%; Score 493.5; DB 15; Length 954;
Best Local Similarity 28.9%; Pred. No. 2.7e-24;
Matches 145; Conservative 82; Mismatches 192; Indels 83; Gaps 16;

QY 21 APSQVLFEEIDYKEIEVEEVGAGVVCCKAKWRAKQVAIKQIESERKAFIV---- 76
Db 83 APAGLQLPQELPFHELOLEELIIGVGKGVTRALWEGEEVAVKARLDPEKDPVTAEOV 142
QY 77 --ELRQLSRVNHNPVIVKLYGACLPN--VCLVMVEAEGGSLYNVLHGAEPPLPYTTAAHAMS 132
Db 143 QCEARLFGALQHPNIIALRGACLPNPHLCLVMVEYARGGALSRLVLAGRRVPHV---LVN 198
QY 133 WCLQCSQGVAYLHSMQKALIHRLDKPPNLLV-----AGGTVLKICDFTGACD-IQT 184
Db 199 WAVQVARGMNYLHNDAPVPIIHRDLKSNIIILAEIENHNLDATVLKITDGLAREWHKT 258
QY 185 HMTNKGSAAMWAVEFEGSNYSKCDVFSNGIILWEVITRRKPFDEIGGPAFRIMWAV- 243
Db 259 TKMSAAGTYAMWAVEVIRLSFSKSSDWSVGVLLWELLTGEVYREI--DALAVAYGVA 316
QY 244 HNGTRPPLIKLPKPIESLMTRCKSKDPSQRPSEEEIVKIMTHL-----MRYFPGA 294
Db 317 MNKLTLPSTCPEPFARLLDEECWDPDHPGRDPFGSILKRLVETQSQALFQMPLESFHS 376
QY 295 DEPLOYPCQY-----SDEGQNSATSTGSGF-----MDTAST 325
Db 377 QEDWKLEIQHMFDDLTTRKELRSEEEILLRAAQEQRFQELRRREQELAREMDI--- 433
QY 326 NTSNKSDDTNWQVPATNDITKRLSKLLKNQAKQSSG---RLSLGASHGSSVESLPPT 382
Db 434 -VERELHLLMCQLSQEKPRVKRKKGNPKRAVLKLRGSSHSILSPSGFEHKITVQASPTL 492
QY 383 SEGKMSADMSEIARIAATTGNGQFRRRSIQDLTVTGTGPGQVSSRSSSPSVRMITTS 442
```

Db 493 D--KRKSGDASPPASPSII-----PRLRAIR--LTPVDCGSSGSSGSGTWSRGG 542

QY 443 PTSEK-----PTRSHWPDDDS 459

Db 543 PPKBELVGGKKKGRTWGSPST 564

RESULT 11

US-10-263-929-187

; Sequence 187, Application US/10263929

; Publication No. US20040067535A1

; GENERAL INFORMATION:

; APPLICANT: Kim, Jaeseob

; APPLICANT: Galant, Ron

; TITLE OF INVENTION: Alzheimer's Disease Linked Genes

; FILE REFERENCE: LSD-07417

; CURRENT APPLICATION NUMBER: US/10/263,929

; CURRENT FILING DATE: 2002-10-03

; NUMBER OF SEQ ID NOS: 213

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 187

; LENGTH: 954

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-263-929-187

Query Match 16.4%; Score 493.5; DB 15; Length 954;

Best Local Similarity 28.9%; Pred. No. 2.7e-24;

Matches 145; Conservative 82; Mismatches 192; Indels 83; Gaps 16;

QY 21 APSQVLFEEIDYKEIEVEVVGARGVGVCKAKWRAKDVAIKQIESERKAFIV---- 76

Db 83 APAGQLQPEIPFHELQLEEEIIGVGFGKVRALWGEVAVKAAARLDPEKDPVATAEQV 142

QY 77 --ELRLSRVNHNPNTVKLYGACLN--VCLVMEYAEAGSLYNVLHGAEPLPYTTAAHAMS 132

Db 143 COEARLFGALQHPNIIALRGACLNPHCLVMEYARGGALSRLVLAGRRVPHV----LVN 198

QY 133 WCLQCSQGVAYLHMQPKALIHRLDKPPNLLV-----AGGTVLKICDFGTACD-IQT 184

Db 199 WAVQVARGMNYLHNDAPVPIIHRLDKSNILILEAENHNLDATVLKITDFGLAREWHKT 258

QY 195 HMTNKGSAAMWAPVFEKSNYSEKDYFSGWIIILWEVITRKRKPDDEIGGPAFRIMWAV- 243

Db 259 TKMSAAGTYAWMAPEVIRLSFSKSDVMSFGVLLWELLTGEVYPREI--DALAVAYGVA 316

QY 244 HNGTRPPLIKNLPKPIESIMTRCWSKDPSPQSMEEIVKIMTHL-----MRYFFGA 294

Db 317 MNKLTLPSTCPEPFARLLEECWDPDPHGRDPDGSILKRLVEIQSALFQMPLESFHS 376

QY 295 DEPLOYQCY-----SDEGQSNSTSTGSP-----MDIATST 325

Db 377 QEDWKLEIQHMFDDLTKEKELRSREELLRAAQEQRFEEQLRREQEQLAEREMDI--- 433

QY 326 NTSNKSNTNMEQVPATNDITKESKLKNQAKQOSESQ--RLSIGASHGSSVESLPT 382

Db 434 -VERELHMLLCQLSQEKFVRKKNFRKRLKUREGSHISLPSGFEHKITVQASPTL 492

QY 383 SEGKMSADMSEIARIAATTNGGQPRRRSIQDLTWTGTEPQCVSSRSSSPSVRMITTS 442

Db 493 D--KEKSGDASPPASPSII-----PRLRAIR--LTPVDCGSSGSSGSGTWSRGG 542

QY 443 PTSEK-----PTRSHWPDDDS 459

Db 543 PPKBELVGGKKKGRTWGSPST 564

RESULT 12

US-10-664-421-136

; Sequence 136, Application US/10664421

; Publication No. US20040142864A1

; GENERAL INFORMATION:

; APPLICANT: BREMER, RYAN

; APPLICANT: IBRAHIM, PRAEHA

; APPLICANT: KOMAR, ABHINAV

; APPLICANT: MANDIYAN, VALSAN

; APPLICANT: MILBURN, MICHAEL V.

; TITLE OF INVENTION: CRYSTAL STRUCTURE OF PIM-1 KINASE

; FILE REFERENCE: 039363/0703

; CURRENT APPLICATION NUMBER: US/10/664,421

; CURRENT FILING DATE: 2003-09-16

; PRIOR APPLICATION NUMBER: 60/412,341

; PRIOR FILING DATE: 2002-09-20

; PRIOR APPLICATION NUMBER: 60/411,398

; PRIOR FILING DATE: 2002-09-16

; NUMBER OF SEQ ID NOS: 169

; SOFTWARE: PatentIn ver. 3.2

; SEQ ID NO 136

; LENGTH: 330

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-664-421-136

Query Match 16.0%; Score 482.5; DB 16; Length 330;

Best Local Similarity 38.4%; Pred. No. 3.9e-24;

Matches 107; Conservative 51; Mismatches 98; Indels 23; Gaps 7;

QY 21 APSQVLFEEIDYKEIEVEVVGARGVGVCKAKWRAKDVAIKQIESERKAFIV---- 76

Db 4 APAGQLQPEIPFHELQLEEEIIGVGFGKVRALWGEVAVKAAARLDPEKDPVATAEQV 63

QY 77 --ELRLSRVNHNPNTVKLYGACLN--VCLVMEYAEAGSLYNVLHGAEPLPYTTAAHAMS 132

Db 64 COEARLFGALQHPNIIALRGACLNPHCLVMEYARGGALSRLVLAGRRVPHV----LVN 119

QY 133 WCLQCSQGVAYLHMQPKALIHRLDKPPNLLV-----AGGTVLKICDFGTACD-IQT 184

Db 120 WAVQVARGMNYLHNDAPVPIIHRLDKSNILILEAENHNLDATVLKITDFGLAREWHKT 179

QY 185 HMTNKGSAAMWAPVFEKSNYSEKDYFSGWIIILWEVITRKRKPDDEIGGPAFRIMWAV- 243

Db 180 TKMSAAGTYAWMAPEVIRLSFSKSDVMSFGVLLWELLTGEVYPREI--DALAVAYGVA 237

QY 244 HNGTRPPLIKNLPKPIESIMTRCWSKDPSPQSMEEIVK 282

Db 238 MNKLTLPSTCPEPFARLLEECWDPDPHGRDPDGSILK 276

RESULT 13

US-10-296-115-837

; Sequence 837, Application US/10296115

; Publication No. US20040053248A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq Inc

; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides

; FILE REFERENCE: 784PCT

; CURRENT APPLICATION NUMBER: US/10/296,115

; CURRENT FILING DATE: 2002-11-18

; PRIOR APPLICATION NUMBER: US09/488,725

; PRIOR FILING DATE: 2000-01-21

; PRIOR APPLICATION NUMBER: US09/552,317

; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 1478

; SEQ ID NO 837

; LENGTH: 473

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-296-115-837

Query Match 16.0%; Score 482; DB 15; Length 473;

Best Local Similarity 30.4%; Pred. No. 6.6e-24;

Matches 134; Conservative 81; Mismatches 156; Indels 70; Gaps 18;

QY 20 EAPSQVLFEEIDYKEIEVEVVGARGVGVCKAKW--RAKDVAIK---QIESERKAF 74

Db 18 EMSSLGASFVQIKFDDLLQFFENCGGSGSVYRAKWISQDKREAVAVKLLKIEKAE---- 73

Qy 75 IVELRQISRVNHPNIVKLYGACLP--VCLWVEYAEAGSLYNVLHG--AEPLPYTAAHA 130  
Db 74 -----ILSVLSHRNIIQFYGVILEPPNPGYIVTEYASLSGLDYINSNRSEEM--DMDHI 125  
Qy 131 MSWCLQCSOGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDFTACDIQTHMTNNK 190  
Db 126 MTWATDVAKGWHYLLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTHMS 183  
Qy 191 --GSAAMWAPVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGT 247  
Db 184 LVGTFPMWAPVIOQLPVSCTDYSGVVLWMLTREVPFKLEG--LQVAVLWVEKNE 241  
Qy 248 RPPLIKLKPDESIMTRCKSKPQSPRSMEIVKIMTHLMRYFPGADEPLOYPCQYSDE 307  
Db 242 RLTFPSGPRFAELLHQWEADAKKRFQIISIL-----E 279  
Qy 308 GQSNATSTGFMFIASNTNSKDTNNEQVPATNDTIKRLSKLLKNQAKQSGESGLS 367  
Db 280 SMSNDT----SLPKCNKSNFLHNKAEWRCE-TEATLERKKLERDLSPKEQBLKERRRLK 334  
Qy 368 L---GASHGVSVELPPTS-----EGKRMADMSIEIARIAATT-GNGQPRRRSIQ 414  
Db 335 MWEQKLEQSNTPLLPLAARMSSESYFESKTESNSAEMSCQITATSNRSGHGMNPSLQ 394  
Qy 415 DLTVTGTEPGQVSRSSSPSV 435  
Db 395 AMMLMGF--GDIFSMNKGAV 413

RESULT 14  
US-09-757-982-5  
; Sequence 5, Application US/09757982  
; Patent No. US20020094559A1  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MH-050  
; CURRENT APPLICATION NUMBER: US/09/757,982  
; PRIOR FILING DATE: 2001-01-10  
; PRIOR APPLICATION NUMBER: 09/163,115  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-757-982-5

Query Match 16.0%; Score 481; DB 9; Length 455;  
Best Local Similarity 30.4%; Pred. No. 7.3e-24;  
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;  
Qy 27 NFEIDYKEIEVEEVGARGVGVCKAKW--RAKDVAIK---QIESESERKAFIVELRQL 81  
Db 7 SFVQIKFDDLOFFENCNGSGSGSVYRAKWI SQDKREVAVKLLKIEKEAE-----IL 57  
Qy 82 SRVNHNPINVLKLYGACLP--VCLWVEYAEAGSLYNVLHG--AEPLPYTAAHMSWCLQC 137  
Db 58 SVLSHRNIIQFYGVILEPPNPGYIVTEYASLSGLDYINSNRSEEM--DMDHIWATDV 114  
Qy 138 SQGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDFTACDIQTHMTNNK--GSAAM 195  
Db 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTPPW 172  
Qy 196 MAPEVFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254  
Db 173 MAPEVIOQLPVSCTDYSGVVLWMLTREVPFKLEG--LQVAVLWVEKNERLTISS 230  
Qy 255 LPKPIESIMTRCKSKPQSPRSMEIVKIMTHLMRYFPGADEPLOYPCQYSDEGOSNAT 314  
Db 231 CPRSFAELLHQWEADAKKRFQIISIL-----E SMSNDT- 267

Qy 315 STGFMFIASNTNSKDTNNEQVPATNDTIKRLSKLLKNQAKQSGESGLSL--GAS 371  
Db 268 ---SLPKCNKSNFLHNKAEWRCE-TEATLERKKLERDLSPKEQBLKERRRLKMWELKLT 323  
Qy 372 HGSVSELSPTS-----EGKRMADMSIEIARIAATT-GNGQPRRRSIQDLTVTGT 421  
Db 324 EQSNTPLLPLAARMSSESYFESKTESNSAEMSCQITATSNRSGHGMNPSLQAMLMGF 383  
Qy 422 EPGQVSRSSSPSV 435  
Db 384 --GDIFSMNKGAV 395

RESULT 15  
US-10-094-749-2477  
; Sequence 2477, Application US/10094749  
; Publication No. US20030219741A1  
; GENERAL INFORMATION:  
; APPLICANT: ISOGAI, TAKAO  
; APPLICANT: SUGIYAMA, TOMOYASU  
; APPLICANT: OTSUKI, TETSUJI  
; APPLICANT: WAKAMATSU, AI  
; APPLICANT: SATO, HIROYUKI  
; APPLICANT: ISHII, SHIZUKO  
; APPLICANT: YAMAMOTO, JUN-ICHI  
; APPLICANT: ISONO, YUUKO  
; APPLICANT: HIO, YURI  
; APPLICANT: OTSUKA, KAORU  
; APPLICANT: NAGAI, KEIICHI  
; APPLICANT: IRIE, RYOTARO  
; APPLICANT: TAMECHIKA, ICHIRO  
; APPLICANT: SEKI, NAOHICO  
; APPLICANT: YOSHIKAWA, TSUTOMU  
; APPLICANT: OTSUKA, MOTOKUKI  
; APPLICANT: NAGAHARI, KENJI  
; APPLICANT: MASUHO, YASUHIKO  
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
; FILE REFERENCE: 084335/0160  
; CURRENT APPLICATION NUMBER: US/10/094,749  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/350,435  
; PRIOR FILING DATE: 2002-01-24  
; PRIOR APPLICATION NUMBER: JP 2001-328381  
; PRIOR FILING DATE: 2001-09-14  
; NUMBER OF SEQ ID NOS: 3381  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2477  
; LENGTH: 455  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-094-749-2477

Query Match 16.0%; Score 481; DB 14; Length 455;  
Best Local Similarity 30.4%; Pred. No. 7.3e-24;  
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;  
Qy 27 NFEIDYKEIEVEEVGARGVGVCKAKW--RAKDVAIK---QIESESERKAFIVELRQL 81  
Db 7 SFVQIKFDDLOFFENCNGSGSGSVYRAKWI SQDKREVAVKLLKIEKEAE-----IL 57  
Qy 82 SRVNHNPINVLKLYGACLP--VCLWVEYAEAGSLYNVLHG--AEPLPYTAAHMSWCLQC 137  
Db 58 SVLSHRNIIQFYGVILEPPNPGYIVTEYASLSGLDYINSNRSEEM--DMDHIWATDV 114  
Qy 138 SQGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDFTACDIQTHMTNNK--GSAAM 195  
Db 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTPPW 172  
Qy 196 MAPEVFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254  
Db 173 MAPEVIOQLPVSCTDYSGVVLWMLTREVPFKLEG--LQVAVLWVEKNERLTISS 230

```

QY 255 LPKPIESLWTRCWSKPSORPSMEEIVKIMTHLMRYFPGADEPLQYPCQYSDGOSNSAT 314
Db 231 CPRSFAELHQWEADAKRPSFKQIISIL-----ESMSNDT- 267
QY 315 STGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLESKILKNQAKQOSESGRLSL---GAS 371
Db 268 ---SLPDKCNSFLHNKAERCE-IEATLERLKKLBERDLSFKQELKERERRLKMWEOKLT 323
QY 372 HGSSVESLPPTS-----EGKRMSADMSEIARIAATT-GNGQPRRSIQDLTVTGT 421
Db 324 EQSNTPLLLPLAARMSEESYFESKTEESNAEMSCQITATSNGEHGGMNPSLQAMLMGF 383
QY 422 EPQVSSRSRSPSV 435
Db 384 --GDIFSMNKAGAV 395

```

Search completed: December 22, 2004, 14:15:13  
Job time : 105.717 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 22, 2004, 13:58:53 ; Search time 25.1302 Seconds  
(without alignments)  
1330.044 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAAPRRSLQEQPSWTDD.....AEFVRLMSVDKGEQSVVTAP 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pap:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pap:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pap:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pap:\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pap:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2580	100.0	504	2	US-08-752-891-2
2	2580	100.0	504	2	US-09-144-178-2
3	2580	100.0	504	3	US-09-466-854-2
4	2580	100.0	504	4	US-09-519-279-2
5	2580	100.0	504	4	US-10-118-895-2
6	2580	100.0	517	4	US-09-519-279-11
7	2580	100.0	517	4	US-10-118-895-11
8	2577	99.9	513	4	US-09-529-279-43
9	2577	99.9	513	4	US-10-158-895-43
10	2575	99.8	504	2	US-08-752-891-6
11	2575	99.8	504	2	US-09-144-178-6
12	2575	99.8	504	3	US-09-406-854-6
13	391	15.2	217	4	US-09-270-767-32865
14	178	6.9	392	3	US-09-013-881-2
15	178	6.9	392	4	US-09-612-473-2
16	178	6.9	392	4	US-09-724-730-2
17	170.5	6.6	372	4	US-09-973-963-4
18	151	5.9	335	4	US-09-270-767-43565
19	150	5.8	669	4	US-09-270-767-46055
20	147	5.7	1512	4	US-09-328-352-5163
21	146	5.7	241	4	US-09-270-767-32790
22	146	5.7	241	4	US-09-270-767-48007
23	131	5.1	454	4	US-09-538-092-1156
24	128	5.0	286	4	US-09-270-767-61602
25	128	5.0	382	4	US-09-538-092-1086
26	125.5	4.9	546	3	US-08-935-855-20
27	125.5	4.9	546	4	US-09-538-092-827

28 124 4.8 306 2 US-08-822-701-8 Sequence 8, Appli  
29 124 4.8 306 3 US-08-935-855-8 Sequence 8, Appli  
30 121 4.7 531 4 US-09-248-796A-15560 Sequence 15560, A  
31 118 4.6 390 4 US-09-206-646-4 Sequence 4, Appli  
32 118 4.6 392 2 US-08-822-701-2 Sequence 2, Appli  
33 118 4.6 392 3 US-08-935-855-22 Sequence 22, Appli  
34 118 4.6 542 3 US-08-935-855-22 Sequence 22, Appli  
35 117.5 4.6 309 3 US-08-822-701-7 Sequence 7, Appli  
36 117.5 4.6 309 3 US-08-935-855-7 Sequence 7, Appli  
37 117.5 4.6 393 4 US-09-538-092-34 Sequence 34, Appli  
38 117 4.5 390 2 US-08-873-093-3 Sequence 3, Appli  
39 117 4.5 390 2 US-09-206-646-3 Sequence 3, Appli  
40 117 4.5 677 1 US-08-188-582-13 Sequence 13, Appli  
41 117 4.5 677 1 US-08-646-715-13 Sequence 13, Appli  
42 117 4.5 695 4 US-09-538-092-1164 Sequence 1164, A  
43 117 4.5 695 4 US-09-248-796A-18020 Sequence 18020, A  
44 116 4.5 314 2 US-08-822-701-10 Sequence 10, Appli  
45 116 4.5 314 3 US-08-935-855-10 Sequence 10, Appli

## ALIGNMENTS

## RESULT 1

US-08-752-891-2  
; Sequence 2, Application US/08752891  
; Patent No. 5837819  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiko  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752.891  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-752-891-2

Query Match 100.0% Score 2580; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 3.5e-235;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQRRLSLOEQPSWTTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
Db 1 MAAQRRLSLOEQPSWTTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
QY 61 NCFLYGVFNGYDGNRVNFAQRLSABELLGQNAEHAADVRRVLLQAFDVVERSFLES 120  
Db 61 NCFLYGVFNGYDGNRVNFAQRLSABELLGQNAEHAADVRRVLLQAFDVVERSFLES 120  
QY 121 IDALA EKASLQSLPEGVPQHQLPPQYQKILERLKTLEIRISGAMAVVAVLLNNKLYV 180  
Db 121 IDALA EKASLQSLPEGVPQHQLPPQYQKILERLKTLEIRISGAMAVVAVLLNNKLYV 180  
QY 181 ANVTGNRALLCKSTVDGLQVLTQNLVNDHTTENEDELFRLSQLGLDAGIKQVGIICGQEST 240  
Db 181 ANVTGNRALLCKSTVDGLQVLTQNLVNDHTTENEDELFRLSQLGLDAGIKQVGIICGQEST 240  
QY 241 RRGDYKVKYGYTDIDLLSAAKSPIIAEPIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300  
Db 241 RRGDYKVKYGYTDIDLLSAAKSPIIAEPIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300  
QY 301 GPGQANOEIAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGRARPCPRHEDMTL 360  
Db 301 GPGQANOEIAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGRARPCPRHEDMTL 360  
QY 361 LVRNFGYPLGEMSOPTSPAPAGRVYPVSPYSSAQSTKTSVTLSLWMPSSOGQWNG 420  
Db 361 LVRNFGYPLGEMSOPTSPAPAGRVYPVSPYSSAQSTKTSVTLSLWMPSSOGQWNG 420  
QY 421 AHSASTLDEATPTLTNQSPTLTLOSTNHTQSSSSSDGGLFRSRPAHSPPGEDGRVEP 480  
Db 421 AHSASTLDEATPTLTNQSPTLTLOSTNHTQSSSSSDGGLFRSRPAHSPPGEDGRVEP 480  
QY 481 YVDFAEFYRLWSVDHGEQSVVUTAP 504  
Db 481 YVDFAEFYRLWSVDHGEQSVVUTAP 504

## RESULT 2

US-09-144-178-2  
; Sequence 2, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/144,178  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/752,891  
; FILING DATE: 20-NOV-1996  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-144-178-2

Query Match 100.0%; Score 2580; DB 2; Length 504;  
Best Local Similarity 100.0%; Pred. No. 3,5e-235;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQRRLSLOEQPSWTTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
Db 1 MAAQRRLSLOEQPSWTTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
QY 61 NCFLYGVFNGYDGNRVNFAQRLSABELLGQNAEHAADVRRVLLQAFDVVERSFLES 120  
Db 61 NCFLYGVFNGYDGNRVNFAQRLSABELLGQNAEHAADVRRVLLQAFDVVERSFLES 120  
QY 121 IDALA EKASLQSLPEGVPQHQLPPQYQKILERLKTLEIRISGAMAVVAVLLNNKLYV 180  
Db 121 IDALA EKASLQSLPEGVPQHQLPPQYQKILERLKTLEIRISGAMAVVAVLLNNKLYV 180  
QY 181 ANVTGNRALLCKSTVDGLQVLTQNLVNDHTTENEDELFRLSQLGLDAGIKQVGIICGQEST 240  
Db 181 ANVTGNRALLCKSTVDGLQVLTQNLVNDHTTENEDELFRLSQLGLDAGIKQVGIICGQEST 240  
QY 241 RRGDYKVKYGYTDIDLLSAAKSPIIAEPIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300  
Db 241 RRGDYKVKYGYTDIDLLSAAKSPIIAEPIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300  
QY 301 GPGQANOEIAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGRARPCPRHEDMTL 360  
Db 301 GPGQANOEIAAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGRARPCPRHEDMTL 360  
QY 361 LVRNFGYPLGEMSOPTSPAPAGRVYPVSPYSSAQSTKTSVTLSLWMPSSOGQWNG 420  
Db 361 LVRNFGYPLGEMSOPTSPAPAGRVYPVSPYSSAQSTKTSVTLSLWMPSSOGQWNG 420  
QY 421 AHSASTLDEATPTLTNQSPTLTLOSTNHTQSSSSSDGGLFRSRPAHSPPGEDGRVEP 480  
Db 421 AHSASTLDEATPTLTNQSPTLTLOSTNHTQSSSSSDGGLFRSRPAHSPPGEDGRVEP 480  
QY 481 YVDFAEFYRLWSVDHGEQSVVUTAP 504  
Db 481 YVDFAEFYRLWSVDHGEQSVVUTAP 504

## RESULT 3

US-09-406-854-2  
; Sequence 2, Application US/09406854  
; Patent No. 6140042  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/406.854  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/752.891  
FILING DATE: 20-NOV-1996  
APPLICATION NUMBER: JP 8-300856  
FILING DATE: 28-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 8-126282  
FILING DATE: 24-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 17981/111  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 504 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-406-854-2

Query Match 100.0%; Score 2580; DB 3; Length 504;  
Best Local Similarity 100.0%; Pred. No. 3.5e-235;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
Db 1 MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
Qy 61 NCFLYGVFNGYDGNRVNFAQRLSAEALLQLNAEHAADVRRVLLQAFDIVERSFLES 120  
Db 61 NCFLYGVFNGYDGNRVNFAQRLSAEALLQLNAEHAADVRRVLLQAFDIVERSFLES 120  
Qy 121 IDDALEAKASLSQQLPEGVPOHLPPOYQKILERLKTLEISGAMAVVAVLNNKLYV 180  
Db 121 IDDALEAKASLSQQLPEGVPOHLPPOYQKILERLKTLEISGAMAVVAVLNNKLYV 180  
Qy 181 ANVTGNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST 240  
Db 181 ANVTGNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST 240  
Qy 241 RRIIDYKVKYGYTTIDLLSAKSPPIIAEPIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300  
Db 241 RRIIDYKVKYGYTTIDLLSAKSPPIIAEPIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300  
Qy 301 GPGQANQEIAMIDTEFAKQTSLSDAVAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360  
Db 301 GPGQANQEIAMIDTEFAKQTSLSDAVAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360  
Qy 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPSVYSSAQSTKSTVTLVLMSEGLYKALEAAH 420  
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPSVYSSAQSTKSTVTLVLMSEGLYKALEAAH 420  
Qy 421 AHSASTLDEATPTLTNQSPFTLTQSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
Db 421 AHSASTLDEATPTLTNQSPFTLTQSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
Qy 481 YVDFAEFRLMSVDHGEQSVVTAP 504  
Db 481 YVDFAEFRLMSVDHGEQSVVTAP 504

RESULT 4

US-09-529-279-2  
Sequence 2, Application US/09529279  
Patent No. 6451617  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES  
FILE REFERENCE: 053466/0278  
CURRENT APPLICATION NUMBER: US/09/529,279  
CURRENT FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: PCT/JP98/04796  
PRIOR FILING DATE: 1998-10-22  
PRIOR APPLICATION NUMBER: JP 9/290188  
PRIOR FILING DATE: 1997-10-22  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 504  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-529-279-2

Query Match 100.0%; Score 2580; DB 4; Length 504;  
Best Local Similarity 100.0%; Pred. No. 3.5e-235;  
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
Db 1 MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
Qy 61 NCFLYGVFNGYDGNRVNFAQRLSAEALLQLNAEHAADVRRVLLQAFDIVERSFLES 120  
Db 61 NCFLYGVFNGYDGNRVNFAQRLSAEALLQLNAEHAADVRRVLLQAFDIVERSFLES 120  
Qy 121 IDDALEAKASLSQQLPEGVPOHLPPOYQKILERLKTLEISGAMAVVAVLNNKLYV 180  
Db 121 IDDALEAKASLSQQLPEGVPOHLPPOYQKILERLKTLEISGAMAVVAVLNNKLYV 180  
Qy 181 ANVTGNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST 240  
Db 181 ANVTGNRALLCKSTVDGLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIICGQEST 240  
Qy 241 RRIIDYKVKYGYTTIDLLSAKSPPIIAEPIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300  
Db 241 RRIIDYKVKYGYTTIDLLSAKSPPIIAEPIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300  
Qy 301 GPGQANQEIAMIDTEFAKQTSLSDAVAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360  
Db 301 GPGQANQEIAMIDTEFAKQTSLSDAVAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360  
Qy 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPSVYSSAQSTKSTVTLVLMSEGLYKALEAAH 420  
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPSVYSSAQSTKSTVTLVLMSEGLYKALEAAH 420  
Qy 421 AHSASTLDEATPTLTNQSPFTLTQSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
Db 421 AHSASTLDEATPTLTNQSPFTLTQSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
Qy 481 YVDFAEFRLMSVDHGEQSVVTAP 504  
Db 481 YVDFAEFRLMSVDHGEQSVVTAP 504

RESULT 5

US-10-158-895-2  
Sequence 2, Application US/10158895  
Patent No. 6551840  
GENERAL INFORMATION:  
APPLICANT: ONO, KOICHIRO  
APPLICANT: OHTOMO, TOSHIHIKO  
APPLICANT: TSUCHIYA, MASAYUKI  
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

```
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR FILING DATE: 2002-06-03
; PRIOR FILING DATE: 2000-04-11
; PRIOR FILING DATE: 2000-04-11
; PRIOR FILING DATE: 1998-10-22
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-2

Query Match      100.0%; Score 2580; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 3.5e-235;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Qy 61 NCFLYGVFNCGYDGNRVNFVAQRLSAELLGQLNAEHAEDVRRVLLQAFDVVERSFLES 120
Db 61 NCFLYGVFNCGYDGNRVNFVAQRLSAELLGQLNAEHAEDVRRVLLQAFDVVERSFLES 120
Qy 121 IDALAABKASLSQSLPEGVPOHQLPPQYQKILERLKTLEIREISGAMAVAVLNNKLYV 180
Db 121 IDALAABKASLSQSLPEGVPOHQLPPQYQKILERLKTLEIREISGAMAVAVLNNKLYV 180
Qy 181 ANVGTNRALLCKSTVDCGLQVTLNVNHTTENEDELFRLSQGLDAGIKQVGIICGQEST 240
Db 181 ANVGTNRALLCKSTVDCGLQVTLNVNHTTENEDELFRLSQGLDAGIKQVGIICGQEST 240
Qy 241 RRGIDYKVKYGYTDIDLLSAAKSPITAEPEIHAQPLDGVTFGLVMSGLYKALEAAH 300
Db 241 RRGIDYKVKYGYTDIDLLSAAKSPITAEPEIHAQPLDGVTFGLVMSGLYKALEAAH 300
Qy 301 GPGQANQEIAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Qy 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSSAQSTKTSVTLSLWMPSSQGMVNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSSAQSTKTSVTLSLWMPSSQGMVNG 420
Qy 421 AHSASTLDEATPTLTNQSPITLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNQSPITLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Qy 481 YVDFAEFYRLMSVDHGEQSVVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVVTAP 504

RESULT 6
US-10-158-895-11
; Sequence 11, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR FILING DATE: 2000-04-11
; PRIOR FILING DATE: 1998-10-22
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-11

Query Match      100.0%; Score 2580; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 3.7e-235;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Qy 61 NCFLYGVFNCGYDGNRVNFVAQRLSAELLGQLNAEHAEDVRRVLLQAFDVVERSFLES 120
Db 61 NCFLYGVFNCGYDGNRVNFVAQRLSAELLGQLNAEHAEDVRRVLLQAFDVVERSFLES 120
Qy 121 IDALAABKASLSQSLPEGVPOHQLPPQYQKILERLKTLEIREISGAMAVAVLNNKLYV 180
Db 121 IDALAABKASLSQSLPEGVPOHQLPPQYQKILERLKTLEIREISGAMAVAVLNNKLYV 180
Qy 181 ANVGTNRALLCKSTVDCGLQVTLNVNHTTENEDELFRLSQGLDAGIKQVGIICGQEST 240
Db 181 ANVGTNRALLCKSTVDCGLQVTLNVNHTTENEDELFRLSQGLDAGIKQVGIICGQEST 240
Qy 241 RRGIDYKVKYGYTDIDLLSAAKSPITAEPEIHAQPLDGVTFGLVMSGLYKALEAAH 300
Db 241 RRGIDYKVKYGYTDIDLLSAAKSPITAEPEIHAQPLDGVTFGLVMSGLYKALEAAH 300
Qy 301 GPGQANQEIAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMIDTEFAKQTSLDVAQAQVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Qy 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSSAQSTKTSVTLSLWMPSSQGMVNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSSAQSTKTSVTLSLWMPSSQGMVNG 420
Qy 421 AHSASTLDEATPTLTNQSPITLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNQSPITLTQSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Qy 481 YVDFAEFYRLMSVDHGEQSVVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVVTAP 504

RESULT 7
US-10-158-895-11
; Sequence 11, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR FILING DATE: 2000-04-11
; PRIOR FILING DATE: 1998-10-22
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-11
```

```
; ORGANISM: Homo sapiens
US-10-158-895-11

Query Match      100.0%; Score 2580; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 3.7e-235;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQRSLQSEQQPSWTDLLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQQPSWTDLLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60

Qy 61 NCFLYGVFNGYDGNRVTFVQRLSAEALLLQGLNAEHAADVRRVLLQAFDVVERSFLS 120
Db 61 NCFLYGVFNGYDGNRVTFVQRLSAEALLLQGLNAEHAADVRRVLLQAFDVVERSFLS 120

Qy 121 IDDAEAKASLQSQLEPGVPQHLPPOYQKILERLKTLEISGGAMAVAVLNNKLYV 180
Db 121 IDDAEAKASLQSQLEPGVPQHLPPOYQKILERLKTLEISGGAMAVAVLNNKLYV 180

Qy 181 ANVTGNRALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGTKQVGIICQSEST 240
Db 181 ANVTGNRALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGTKQVGIICQSEST 240

Qy 241 RRIQDYKVYGYTDDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300
Db 241 RRIQDYKVYGYTDDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300

Qy 301 GPGQANQEIAMI DTEFAKQTSLSDAVAQVVDVRVKRIHSDFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMI DTEFAKQTSLSDAVAQVVDVRVKRIHSDFASGGERARFCPRHEDMTL 360

Qy 361 LVRNFGYPLGEMSQTPSPAPAGGRVVPVSPYSSAOSTSKTSVTLNVPSCQMVNG 420
Db 361 LVRNFGYPLGEMSQTPSPAPAGGRVVPVSPYSSAOSTSKTSVTLNVPSCQMVNG 420

Qy 421 AHSASTLDEAFTLTNQSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEAFTLTNQSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480

Qy 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVTAP 504

RESULT 8
US-09-529-279-43
; Sequence 43, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-43

Query Match      99.9%; Score 2577; DB 4; Length 513;
Best Local Similarity 99.8%; Pred. No. 7e-235;
Matches 503; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQRSLQSEQQPSWTDLLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQQPSWTDLLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60

Qy 61 NCFLYGVFNGYDGNRVTFVQRLSAEALLLQGLNAEHAADVRRVLLQAFDVVERSFLS 120
Db 61 NCFLYGVFNGYDGNRVTFVQRLSAEALLLQGLNAEHAADVRRVLLQAFDVVERSFLS 120

Qy 121 IDDAEAKASLQSQLEPGVPQHLPPOYQKILERLKTLEISGGAMAVAVLNNKLYV 180
Db 121 IDDAEAKASLQSQLEPGVPQHLPPOYQKILERLKTLEISGGAMAVAVLNNKLYV 180

Qy 181 ANVTGNRALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGTKQVGIICQSEST 240
Db 181 ANVTGNRALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGTKQVGIICQSEST 240

Qy 241 RRIQDYKVYGYTDDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300
Db 241 RRIQDYKVYGYTDDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300

Qy 301 GPGQANQEIAMI DTEFAKQTSLSDAVAQVVDVRVKRIHSDFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMI DTEFAKQTSLSDAVAQVVDVRVKRIHSDFASGGERARFCPRHEDMTL 360

Qy 361 LVRNFGYPLGEMSQTPSPAPAGGRVVPVSPYSSAOSTSKTSVTLNVPSCQMVNG 420
Db 361 LVRNFGYPLGEMSQTPSPAPAGGRVVPVSPYSSAOSTSKTSVTLNVPSCQMVNG 420

Qy 421 AHSASTLDEAFTLTNQSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEAFTLTNQSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480

Qy 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVTAP 504

RESULT 9
US-10-158-895-43
; Sequence 43, Application US/10158895
; Patent No. 6551840
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-43

Query Match      99.9%; Score 2577; DB 4; Length 513;
Best Local Similarity 99.8%; Pred. No. 7e-235;
Matches 503; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAAQRSLQSEQQPSWTDLLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQQPSWTDLLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60

Qy 61 NCFLYGVFNGYDGNRVTFVQRLSAEALLLQGLNAEHAADVRRVLLQAFDVVERSFLS 120
Db 61 NCFLYGVFNGYDGNRVTFVQRLSAEALLLQGLNAEHAADVRRVLLQAFDVVERSFLS 120

Qy 121 IDDAEAKASLQSQLEPGVPQHLPPOYQKILERLKTLEISGGAMAVAVLNNKLYV 180
Db 121 IDDAEAKASLQSQLEPGVPQHLPPOYQKILERLKTLEISGGAMAVAVLNNKLYV 180

Qy 181 ANVTGNRALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGTKQVGIICQSEST 240
Db 181 ANVTGNRALLCKSTVDGLQVTLNVDHTTENEDELFRLSQLGLDAGTKQVGIICQSEST 240

Qy 241 RRIQDYKVYGYTDDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300
Db 241 RRIQDYKVYGYTDDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300

Qy 301 GPGQANQEIAMI DTEFAKQTSLSDAVAQVVDVRVKRIHSDFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMI DTEFAKQTSLSDAVAQVVDVRVKRIHSDFASGGERARFCPRHEDMTL 360

Qy 361 LVRNFGYPLGEMSQTPSPAPAGGRVVPVSPYSSAOSTSKTSVTLNVPSCQMVNG 420
Db 361 LVRNFGYPLGEMSQTPSPAPAGGRVVPVSPYSSAOSTSKTSVTLNVPSCQMVNG 420

Qy 421 AHSASTLDEAFTLTNQSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEAFTLTNQSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480

Qy 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
```

QY 121 IDDALEKASLSQSLPGVFOHQPLPPQYQKILERLKTLEIRISGAMAVAVLNNKLYV 180  
DB 130 IDDALEKASLSQSLPGVFOHQPLPPQYQKILERLKTLEIRISGAMAVAVLNNKLYV 189  
QY 181 ANVTNRALLCKSVTDGLQVTLQNVHTTENEDELFRLSQGLDAGIKQVGIICGQEST 240  
DB 190 ANVTNRALLCKSVTDGLQVTLQNVHTTENEDELFRLSQGLDAGIKQVGIICGQEST 249  
QY 241 RRGIDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300  
DB 250 RRGIDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 309  
QY 301 GPGQANOEIAAMIDTEFAKQTSLSDAVAQAVVDRVKRIHSDFASGGERARFCPRHEDMTL 360  
DB 310 GPGQANOEIAAMIDTEFAKQTSLSDAVAQAVVDRVKRIHSDFASGGERARFCPRHEDMTL 369  
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSSAQSTSKTSVTLSLWPSQGMVNG 420  
DB 370 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSSAQSTSKTSVTLSLWPSQGMVNG 429  
QY 421 AHSASTLDEATPTLTNOSPILTLQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
DB 430 AHSASTLDEATPTLTNOSPILTLQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 489  
QY 481 YVDFAEFYRLWSVDHGQSVVVTAP 504  
DB 490 YVDFAEFYRLWSVDHGQSVVVTAP 513

RESULT 10  
US-08-752-891-6  
; Sequence 6, Application US/08752891  
; Patent No. 5837819  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/752,891  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-300856  
; FILING DATE: 28-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-126282  
; FILING DATE: 24-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17981/111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 504 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-752-891-6

Query Match 99.8%; Score 2575; DB 2; Length 504;  
Best Local Similarity 99.8%; Pred. No. 1.1e-234;  
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQRRLSLLSQEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
DB 1 MAQRRLSLLSQEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
QY 61 NCFLYGVFNQYDGNRVTFVAQRLSAELLIGQLNAEHAEDVRRVLLQAFDIVERSFLES 120  
DB 61 NCFLYGVFNQYDGNRVTFVAQRLSAELLIGQLNAEHAEDVRRVLLQAFDIVERSFLES 120  
QY 121 IDDALEKASLSQSLPGVFOHQPLPPQYQKILERLKTLEIRISGAMAVAVLNNKLYV 180  
DB 121 IDDALEKASLSQSLPGVFOHQPLPPQYQKILERLKTLEIRISGAMAVAVLNNKLYV 180  
QY 181 ANVTNRALLCKSVTDGLQVTLQNVHTTENEDELFRLSQGLDAGIKQVGIICGQEST 240  
DB 181 ANVTNRALLCKSVTDGLQVTLQNVHTTENEDELFRLSQGLDAGIKQVGIICGQEST 240  
QY 241 RRGIDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300  
DB 241 RRGIDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVLMSEGLYKALEAAH 300  
QY 301 GPGQANOEIAAMIDTEFAKQTSLSDAVAQAVVDRVKRIHSDFASGGERARFCPRHEDMTL 360  
DB 301 GPGQANOEIAAMIDTEFAKQTSLSDAVAQAVVDRVKRIHSDFASGGERARFCPRHEDMTL 360  
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSSAQSTSKTSVTLSLWPSQGMVNG 420  
DB 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSSAQSTSKTSVTLSLWPSQGMVNG 420  
QY 421 AHSASTLDEATPTLTNOSPILTLQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
DB 421 AHSASTLDEATPTLTNOSPILTLQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480  
QY 481 YVDFAEFYRLWSVDHGQSVVVTAP 504  
DB 481 YVDFAEFYRLWSVDHGQSVVVTAP 504

RESULT 11  
US-09-144-178-6  
; Sequence 6, Application US/09144178  
; Patent No. 5989862  
; GENERAL INFORMATION:  
; APPLICANT: MATSUMOTO, Kunihiro  
; APPLICANT: NISHIDA, Eisuke  
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/144,178  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/752,891  
; FILING DATE: 20-NOV-1996

; APPLICATION NUMBER: JP 8-300856  
 ; FILING DATE: 28-OCT-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 8-126282  
 ; FILING DATE: 24-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 17981/111  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202)672-5300  
 ; TELEFAX: (202)672-5399  
 ; TELEX: 904136  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 504 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-144-178-6

Query Match 99.8%; Score 2575; DB 2; Length 504;  
 Best Local Similarity 99.8%; Pred. No. 1.1e-234;  
 Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 MAAQRSLLOSQSQSWTDDLPCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
 Db 1 MAAQRSLLOSQSQSWTDDLPCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
  
 QY 61 NCFLYGVNGYDGNRVNFFVAQRLSAELLGQNAEHAADVRRVLLQAFDVVVERSFLES 120  
 Db 61 NCFLYGVNGYDGNRVNFFVAQRLSAELLGQNAEHAADVRRVLLQAFDVVVERSFLES 120  
  
 QY 121 IDDAEAKASQSQPEGVQHPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180  
 Db 121 IDDAEAKASQSQPEGVQHPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180  
  
 QY 181 ANVGNTRALLCKSTVDGLQVTLNVDDHTTENEDELFRLSQLGLDAGKIKQVGIICQEST 240  
 Db 181 ANVGNTRALLCKSTVDGLQVTLNVDDHTTENEDELFRLSQLGLDAGKIKQVGIICQEST 240  
  
 QY 241 RRIQYKVKYGYTDIDLLSAAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300  
 Db 241 RRIQYKVKYGYTDIDLLSAAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300  
  
 QY 301 GPCQANQEIATAMIDTEFAKQTSLDAVAQAVDVRVKRIHSDTTFASGGERARFCPRHEDMTL 360  
 Db 301 GPCQANQEIATAMIDTEFAKQTSLDAVAQAVDVRVKRIHSDTTFASGGERARFCPRHEDMTL 360  
  
 QY 361 LVNFGYPLGEMSQPTSPAPAGGRVYPVSPYSSAQSTKSTVTLVMPSSQGMVNG 420  
 Db 361 LVNFGYPLGEMSQPTSPAPAGGRVYPVSPYSSAQSTKSTVTLVMPSSQGMVNG 420  
  
 QY 421 AHSASTLDEATPTLTNQSTPTLTQSTNTHTQSSSSSDGGLFRSPAHSLPPGEDGRVEP 480  
 Db 421 AHSASTLDEATPTLTNQSTPTLTQSTNTHTQSSSSSDGGLFRSPAHSLPPGEDGRVEP 480  
  
 QY 481 YVDFAEFYRLWSVDHGEQSVWTAP 504  
 Db 481 YVDFAEFYRLWSVDHGEQSVWTAP 504

RESULT 12  
 US-09-406-854-6  
 ; Sequence 6, Application US/09406854  
 ; Patent No. 6140042  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MATSUMOTO, Kunihiro  
 ; APPLICANT: NISHIDA, Eisuke  
 ; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/406,854  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/752,891  
 ; FILING DATE: 20-NOV-1996  
 ; APPLICATION NUMBER: JP 8-300856  
 ; FILING DATE: 28-OCT-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 8-126282  
 ; FILING DATE: 24-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 17981/111  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202)672-5300  
 ; TELEFAX: (202)672-5399  
 ; TELEX: 904136  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 504 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-406-854-6

Query Match 99.8%; Score 2575; DB 3; Length 504;  
 Best Local Similarity 99.8%; Pred. No. 1.1e-234;  
 Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 1 MAAQRSLLOSQSQSWTDDLPCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
 Db 1 MAAQRSLLOSQSQSWTDDLPCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60  
  
 QY 61 NCFLYGVNGYDGNRVNFFVAQRLSAELLGQNAEHAADVRRVLLQAFDVVVERSFLES 120  
 Db 61 NCFLYGVNGYDGNRVNFFVAQRLSAELLGQNAEHAADVRRVLLQAFDVVVERSFLES 120  
  
 QY 121 IDDAEAKASQSQPEGVQHPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180  
 Db 121 IDDAEAKASQSQPEGVQHPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180  
  
 QY 181 ANVGNTRALLCKSTVDGLQVTLNVDDHTTENEDELFRLSQLGLDAGKIKQVGIICQEST 240  
 Db 181 ANVGNTRALLCKSTVDGLQVTLNVDDHTTENEDELFRLSQLGLDAGKIKQVGIICQEST 240  
  
 QY 241 RRIQYKVKYGYTDIDLLSAAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300  
 Db 241 RRIQYKVKYGYTDIDLLSAAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300  
  
 QY 301 GPCQANQEIATAMIDTEFAKQTSLDAVAQAVDVRVKRIHSDTTFASGGERARFCPRHEDMTL 360  
 Db 301 GPCQANQEIATAMIDTEFAKQTSLDAVAQAVDVRVKRIHSDTTFASGGERARFCPRHEDMTL 360  
  
 QY 361 LVNFGYPLGEMSQPTSPAPAGGRVYPVSPYSSAQSTKSTVTLVMPSSQGMVNG 420  
 Db 361 LVNFGYPLGEMSQPTSPAPAGGRVYPVSPYSSAQSTKSTVTLVMPSSQGMVNG 420  
  
 QY 421 AHSASTLDEATPTLTNQSTPTLTQSTNTHTQSSSSSDGGLFRSPAHSLPPGEDGRVEP 480  
 Db 421 AHSASTLDEATPTLTNQSTPTLTQSTNTHTQSSSSSDGGLFRSPAHSLPPGEDGRVEP 480

```

QY 481 YVDFAEYRLWSVDHGQSVVTP 504
Db 481 YVDFAEYRLWSVDHGQSVVTP 504

RESULT 13
US-09-270-767-32865
; Sequence 32865, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-034
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32865
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32865

Query Match 15.2%; Score 391; DB 4; Length 217;
Best Local Similarity 45.7%; Pred. No. 1.1e-28;
Matches 86; Conservative 36; Mismatches 64; Indels 2; Gaps 2;

QY 151 ILERLKLTERETSGGAMAVAVLLNNKLYVANVGTNRALLCKSTVDG-LQVTLNVDDHT 209
Db 14 ILARKEIDVHLSGGAIVIALVHANKLVFAHVGDTALLCFDDNVLVQVLVDHSL 73

QY 210 ENEBELFRLSQGLDAGTKVGTICGGESTRIGDYKVGYTIDILLSAKSKPIIAE 269
Db 74 NNEDELLRLQQLGLDVKNLRNAQYLGNQTGTRCLGNLYLVKGLYKAPFIIISAAVSBFVIAA 133

QY 270 PEIHAQPLDGVTFGLVLMSEGLYKALEAAHGP-GQANOEIAAMIDTEFAKOTSLDVAQ 328
Db 134 PEIHGFILLDES CRFLVLVSAGYKRIQESKGSYEQTNKQLACLIVENFRKQTDPRMVSQ 193

QY 329 AVVDVRVK 336
Db 194 AVLEIEIQ 201

RESULT 14
US-09-013-881-2
; Sequence 2, Application US/09013881
; Patent No. 6132964
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ For Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,881
; FILING DATE: HERewith

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, LUCY J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0470 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: KIDNNOT02
; CLONE: 195647
US-09-013-881-2

Query Match 6.9%; Score 178; DB 3; Length 392;
Best Local Similarity 22.9%; Pred. No. 4.4e-08;
Matches 83; Conservative 65; Mismatches 116; Indels 98; Gaps 18;

QY 3 AQRRLLOEQPSWTDL---PLCHLSGVGSASNRYSADCKG-----T 44
Db 77 AKRKT---SEBEKNGSEBELVKVKCKASSV-IFGLKGYVAERKGEREQMDAHVILNDIT 132

QY 45 ESHPPEDSWLKRSENNCFLYGVNGYDGNRVTVNAQRLSAELL--LQLMAEHAADV 102
Db 133 EECRPSSLI-----TRVSYFAVFDGHHGIRASKFAAQNHLQNLIRKPKGVISVEKTV 187

QY 103 RVLVLAQFVWVERSEFLESIDDAEAKASLQSLPEGVFQHPQVQPKILERLTEREI 162
Db 188 KECILDITFKHTDEEL-----KQASSQK-----PAWK----- 214

QY 163 SGGAHVAVLLNNKLYVANVGTNRALLCKSTVDGLQ--VTOLNVDDHTTENEDELRLSQ 220
Db 215 -DGSTATCVLAVDNILYIANLGDRAILCRVNEESQKHAALSLSKEHNPTQVEERNRIQK 273

QY 221 LGLDAGKIKQVGIIICGQESTRIGD--YKVKGVTIDILLSAKSKPIIAEPIHGAQ-- 276
Db 274 AG---GNVRDGRVLGVLEVSRSIGDQYK-RGVTSV-----PDIRECQLT 315

QY 277 PLDGVTFGLVLMSEGLYKALEAAHGPQGANQBIAMIDTEFAK-----QTSLDVAQAVVD 332
Db 316 PND---RFILLACDGLFKVFT----PEEAVNFILSCLEDEKIQTRGKSAADARYEAACN 368

QY 333 RV 334
Db 369 RL 370

RESULT 15
US-09-612-473-2
; Sequence 2, Application US/09612473
; Patent No. 6518029
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
; FILE REFERENCE: PF-0470-1 CIP
; CURRENT APPLICATION NUMBER: US/09/612,473
; CURRENT FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 09/013,881
; PRIOR FILING DATE: 1998-02-06

```





This Page Blank (uspto)